

TEXTE

08/2026

Abschlussbericht

Anpassung der Berichtsprozesse für Wasser und Boden

Entwicklung einer übergreifenden und integrierten Infrastruktur für Zustandsdaten im Boden- und Gewässerschutz

von:

Matthias Lüttgert

ENDA GmbH & Co. KG, Berlin

Herausgeber:

Umweltbundesamt

TEXTE 08/2026

REFOPLAN des Bundesministeriums Umwelt,
Naturschutz, nukleare Sicherheit und Verbraucherschutz

Forschungskennzahl 3720 12 202 0
FB001856

Abschlussbericht

Anpassung der Berichtsprozesse für Wasser und Boden

Entwicklung einer übergreifenden und integrierten
Infrastruktur für Zustandsdaten im Boden- und
Gewässerschutz

von

Matthias Lüttgert
ENDA GmbH & Co. KG, Berlin

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

Durchführung der Studie:

ENDA GmbH & Co. KG
Flottwellstr. 15
10785 Berlin

Abschlussdatum:

November 2024

Redaktion:

Fachgebiet II 2.1 Übergreifende Angelegenheiten
Falk Hilliges

Fachgebiet II 2.7 Bodenzustand, Bodenmonitoring
Dr. Marc Marx

DOI:

<https://doi.org/10.60810/openumwelt-8007>

ISSN 1862-4804

Dessau-Roßlau, Januar 2026

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

Kurzbeschreibung: Anpassung der Berichtsprozesse für Wasser und Boden

Das Forschungsprojekt „Anpassung der Berichtsprozesse für Wasser und Boden“ (FKZ 3720122020) zielte darauf ab, die Berichtsprozesse für Fließgewässer, Seen, Grundwasser und Boden beim Umweltbundesamt (UBA) zu vereinheitlichen. Die bisherigen heterogenen Datenformate und Prozesse erschwerten eine effiziente Datenverarbeitung und die Erfüllung von Berichtspflichten. Das entwickelte Fachinformationssystem FIS Wasser-Boden schafft eine einheitliche Datenstruktur und digitale Prozesse von der Datenerfassung bis zur Berichterstellung. Weitere Ziele waren die Förderung von Interoperabilität, OpenData, Digitalisierung und die Flexibilität zur Integration weiterer Umweltmedien wie Meeresumweltdaten (MUDAB).

Das Projekt wurde in fünf Arbeitspakete gegliedert: Erarbeitung eines Fachkonzeptes, technische Umsetzung, Prüfung der Integration weiterer Medien, Bearbeitung von Forschungsfragen und Abschlussdokumentation. Agile Methoden (Scrum) und die Einhaltung der Architekturrichtlinie IT-Bund sicherten eine flexible und standardkonforme Entwicklung.

Das FIS Wasser-Boden ermöglicht eine einheitliche Datenverarbeitung, reduziert Komplexität und nutzt Synergieeffekte durch gemeinsame Werkzeuge wie Abfragewerkzeug und Data Warehouse. Moderne REST-Schnittstellen und Formate wie .xlsx, CSV und JSON fördern Interoperabilität, während automatisierte Datenverarbeitung und geringe Netzwerkanforderungen die Digitalisierung unterstützen. Die Webanwendung integriert Module für Datenanzeige, -bearbeitung, -aggregation und Berichterstellung, z. B. für WISE-6 und GEMS/Water. Qualitätssicherung gewährleistet Datenintegrität.

Das FIS bietet eine skalierbare Basis für zukünftige Umweltüberwachung, mit Potenzial für KI-Nutzung und erweiterte Webdienste. Das Projekt unterstreicht, wie moderne IT-Lösungen die Umweltverwaltung stärken können.

Abstract: Adaptation of Reporting Processes for Water and Soil

The research project “Adaptation of Reporting Processes for Water and Soil” (FKZ 3720122020) aimed to standardise reporting processes for rivers, lakes, groundwater, and soil at the German Environment Agency (UBA). Heterogeneous data formats and processes previously hindered efficient data processing and compliance with reporting obligations. The developed Fachinformationssystem (FIS Water-Soil) establishes a unified data structure and digital workflows from data collection to reporting. Additional goals included enhancing interoperability, promoting OpenData, supporting digitalisation, and ensuring flexibility for integrating further environmental media, such as marine data (MUDAB).

Structured into five work packages—conceptual framework, technical implementation, assessment of additional media integration, research questions, and final documentation—the project employed agile methods (Scrum) and adhered to the Federal IT Architecture Guideline for a flexible, standards-compliant solution.

The FIS Water-Soil enables consistent data processing, reduces complexity, and leverages synergies through shared tools like a query tool and data warehouse. Modern REST interfaces and formats like .xlsx, CSV, and JSON enhance interoperability, while automated processing and low network requirements support digitalisation. The web application includes modules for data display, editing, aggregation, and reporting, e.g., for WISE-6 and GEMS/Water. Quality assurance ensures data integrity.

The FIS provides a scalable foundation for future environmental monitoring, with potential for AI applications and enhanced web services. The project highlights how modern IT solutions can strengthen environmental administration.

Inhaltsverzeichnis

Abbildungsverzeichnis.....	13
Tabellenverzeichnis	13
Abkürzungsverzeichnis	14
Zusammenfassung.....	16
Summary	20
1 Einleitung.....	24
2 Mit der Leistungsbeschreibung verfolgte Ziele und umzusetzende Aufgaben	25
2.1 Arbeitspaket 1 - Erarbeitung eines Fachkonzeptes	25
2.2 Arbeitspaket 2 - Umsetzung des Konzeptes	25
2.3 Arbeitspaket 3 - Prüfung der Einbindung weiterer Umweltmedien	26
2.4 Arbeitspaket 4 - Integrierte Forschungsaspekte und Umsetzung der Forschungsfragen	26
2.5 Arbeitspaket 5 - Abschlussdokumentation	26
3 Übergeordnete Ziele	27
3.1 Interoperabilität.....	28
3.2 OpenData	28
3.3 Schnittstellen	29
3.4 Digitalisierung	29
3.4.1 Ableitung der Kriterien zur Digitalisierung	30
3.5 Synergieeffekte	32
3.6 Agile Methoden	33
3.7 Multikulturelles Team.....	34
3.8 Vorteile in der Zusammenarbeit mit verschiedenen Verwaltungen	34
3.9 Ablösung veralteter Standards	35
3.10 Flexibilität auch für Integration neuer Umweltfelder	35
3.11 Integrative Betrachtung der Umwelt	36
3.12 Unterstützung der Maßnahmenableitung.....	36
3.13 Webdienste.....	36
4 Anforderungen an Konzept und Umsetzung des FIS Wasser-Boden	38
4.1 Technische Umsetzung und Integration in die IT-Umgebung des UBA	38
4.2 Module der Anwendung	42
5 Aufnahme der Nutzungsprozesse	43
5.1 Fließgewässer.....	43
5.2 Seen und Grundwasser	45

5.3	Boden	45
6	Zielprozesse und Übergangsprozess Datenlieferung	46
6.1	Zielprozess Datenlieferung	46
6.2	Zielprozess Produkterstellung.....	50
6.3	Übergangsprozess Datenlieferung.....	52
6.4	Alternativer Zielprozess zur Datenakquisition	52
7	Konzept Anwendung	54
7.1	Anzeige von Daten	54
7.2	Funktionalität zur Nutzung von Parametergruppen.....	55
7.3	Funktionalität zur Nutzung von Messnetzen	55
7.4	Fließgewässer Datenzugriff.....	56
7.5	Seen – Modell und Anwendungsnutzung	56
7.6	Grundwasser	57
7.7	Boden	57
7.8	Prüfungen.....	57
7.9	Editieren von Daten	58
7.10	Abfragewerkzeug	58
7.11	Aggregation von Daten	59
7.12	Reporterstellung EUA/Nitrat-RL.....	60
8	Rollen und Rechte WaBo-Intranet	61
8.1	Rollen von Benutzer:innen.....	61
8.2	Benutzer:innen-Rechte	61
9	Rollen und Rechte WaBo-DMZ.....	65
9.1	Rollen von Benutzer:innen.....	65
9.2	Rollen von Organisationen.....	65
9.3	Benutzer:innen-Rechte	65
10	Datenübertragung und Modellkonsolidierung.....	69
10.1	Modellpotenzial	69
11	Datenmodelle	72
11.1	Datenmodell FIS Wasser-Boden DMZ (WaBo-DMZ)	72
11.2	Datenmodell FIS Wasser-Boden (WaBo-Intranet)	78
11.2.1	Erläuterungen	89
12	Nachrichtenmodelle	91
12.1	Fließgewässer.....	91

12.2	Biologische Parameter	93
12.2.1	Taxonomie und Abundanz	93
12.2.2	Makrophytenausdehnung (max. Tiefe)	93
12.2.3	Chla- und Phaeophytindaten (Konzentrationen) (Zellpigment)	93
12.2.4	EQRs und Messungen (bspw. Schwermetalle, org. Schadstoffe) für die Kompartimente Sediment, Schwebstoff und diverse Organismen	94
12.2.5	Zustandsklassen von 1 bis 5 und chemischer Zustand gut / nicht gut	94
12.3	Seen.....	94
12.4	Grundwasser	95
12.5	Gemeinsames Nachrichtenmodell Wasser	96
12.6	Boden	98
12.7	Tabellenformate.....	99
13	Data Warehouse.....	100
13.1	Befüllung des Data Warehouse und Zugriff auf die Inhalte.....	100
14	Prüfung der Einbindung weiterer Umweltmedien	102
14.1	Meeresumweltdatenbank (MUDAB)	102
14.1.1	Zusammenfassung	102
14.1.2	Arbeitspakete zur Integration der MUDAB in FIS Wasser-Boden.....	103
14.1.3	Detaillierung der Arbeitsschritte	104
14.1.3.1	Importschnittstellen	104
14.1.3.2	Aufarbeitung aller MUDAB-Produkte aus dem Bereich Datenlieferungen	105
14.1.3.3	MUDAB Webpräsenz	106
14.2	Pflanzenschutzmittel (PSM)	106
14.2.1	Tätigkeitsspektren im Bereich PSM	107
14.2.2	Herausforderungen durch PSM	107
15	Integrierte Forschungsaspekte und Umsetzung der Forschungsfragen	108
15.1	Nutzung fremder Webdienste	108
15.2	Ansatz zur medienübergreifenden Sicht und Recherche	108
16	Ausblick	110
17	Quellenverzeichnis	111
A	Standardisierte Erhebungsstruktur zur Prozess- und Anforderungsanalyse	112
A.1	Status quo: Fachdatenbanken und Software.....	112
A.2	Datenarten	112
A.3	Datenmenge.....	113

A.4	Datenimport.....	113
A.4.1	Quelle.....	113
A.4.2	Übermittlungsmodus	113
A.4.3	Import in die Fachdatenbank.....	114
A.4.4	Qualitätssicherung.....	114
A.4.5	Datenkorrektur und -aktualisierung	115
A.5	Manuelle Datenänderungen	115
A.6	Datenanalyse.....	115
A.6.1	Methodik.....	115
A.6.2	Berechnungen	116
A.7	Datenexport	116
A.8	Berichterstattung	116
A.9	Verschiedenes.....	117
B	Prozessanalyse Fließgewässer – Ergebnisse.....	118
B.1	Status quo: Fachdatenbanken und Software.....	118
B.2	Datenarten	118
B.2.1	Stammdaten.....	118
B.2.2	Referenzdaten.....	119
B.2.3	Bewegungsdaten	119
B.2.4	Metadaten	120
B.3	Datenmenge.....	120
B.3.1	Stammdaten.....	120
B.3.2	Referenzdaten.....	121
B.3.3	Bewegungsdaten	121
B.4	Datenimport.....	121
B.4.1	Datenimport Messstellen	121
B.4.2	Messnetze	122
B.4.3	Einheiten	122
B.4.4	Einheitenumrechnungsvorschriften	123
B.4.5	Datenimport Stoffe/Parameter	123
B.4.6	Datenimport Stoffgruppen	124
B.4.7	Datenimport Probenahmematrizen	124
B.4.8	Datenimport Probenahmearten	124
B.4.9	Datenimport Messarten	124

B.4.10	Datenimport Güteklassen	124
B.4.11	Datenimport Messwerte.....	124
B.5	Datenänderungen	125
B.5.1	Aggregierte Messwerte.....	126
B.6	Datenanalyse.....	126
B.6.1	Methodik.....	126
B.6.2	Berechnungen	126
B.7	Datenexport	126
B.8	Berichterstattung	126
B.9	Verschiedenes	126
B.9.1	Unnötige Funktionen oder Daten	127
B.9.2	Fehlende Funktionen oder Daten	127
C	Prozessanalyse Boden	128
C.1	Status quo: Fachdatenbanken und Software.....	128
C.2	Datenarten	128
C.3	Datenmenge.....	129
C.4	Datenimport.....	129
C.4.1	Quelle	129
C.4.2	Übermittlungsmodus	130
C.4.3	Import in die Fachdatenbank.....	130
C.4.4	Qualitätssicherung	131
C.4.5	Datenkorrektur und -aktualisierung	131
C.5	Manuelle Datenänderungen	132
C.6	Datenanalyse.....	132
C.6.1	Methodik.....	132
C.6.2	Berechnungen	133
C.7	Datenexport	133
C.8	Berichterstattung	133
C.9	Verschiedenes	134
D	Prozessanalyse Grundwasser	135
D.1	Status quo: Fachdatenbanken und Software.....	135
D.2	Datenarten	135
D.3	Datenmenge.....	137
D.4	Datenimport.....	138

D.4.1	Datenimport Messstellen	138
D.4.2	Datenimport Messnetze	139
D.4.3	Datenimport Grundwasserkörper	139
D.4.4	Datenimport Parameter.....	139
D.4.5	Datenimport Parametergruppen	140
D.4.6	Datenimport physikalische Größen & Einheiten	140
D.4.7	Datenimport Plausibilität- und Gütegrenzen.....	140
D.4.8	Datenimport Messwerte.....	141
D.5	Manuelle Datenänderungen	142
D.6	Datenanalyse.....	143
D.6.1	Methodik.....	143
D.6.2	Berechnungen	143
D.7	Datenexport	143
D.8	Berichterstattung	144
E	Prozessanalyse Seen.....	145
E.1	Status quo: Fachdatenbanken und Software.....	145
E.2	Datenarten	145
E.2.1	Stammdaten.....	145
E.2.2	Referenzdaten.....	145
E.2.3	Bewegungsdaten	146
E.2.4	Metadaten	146
E.3	Datenmenge.....	147
E.3.1	Stammdaten.....	147
E.3.2	Referenzdaten.....	147
E.3.3	Bewegungsdaten	147
E.4	Datenimport.....	147
E.4.1	Datenimport Messstellen	147
E.4.2	Datenimport Messnetze	148
E.4.3	Datenimport Seen.....	149
E.4.4	Datenimport Flussgebietseinheiten.....	149
E.4.5	Datenimport Parameter.....	150
E.4.6	Datenimport Parametersätze	150
E.4.7	Datenimport physikalische Größen & Einheiten	150
E.4.8	Datenimport Plausibilität- und Gütegrenzen.....	151

E.4.9	Datenimport Messwerte	151
E.5	Manuelle Datenänderungen	152
E.6	Datenanalyse.....	153
E.6.1	6.1 Methodik.....	153
E.6.2	6.2 Berechnungen	153
E.7	Datenexport	153
E.8	Berichterstattung	154
F	MUDAB Bestandsaufnahme	155
F.1	Auswertung bereitgestellter Materialien	155
F.2	Oracle-Datenbank MUDAB	155
F.2.1	Messwerte	156
F.2.2	Parameter	157
F.3	Templates.....	158
F.4	Webpräsenz und deren Berichtsformate	159
F.5	MUDAB Datenimporte	160
F.6	Außendarstellung MUDAB derzeit.....	162
F.7	Untersuchungen mit der Zielstellung einer Datenübernahme	162

Abbildungsverzeichnis

Abbildung 1:	FIS Wasser-Boden IT-Architektur	41
Abbildung 2:	Module der Anwendung.....	42
Abbildung 3:	Ist-Prozess Datenlieferung Fließgewässer	44
Abbildung 4:	Zielprozess Datenlieferung (strategische Sicht)	47
Abbildung 5:	Zielprozess Datenlieferung (Gesamtprozess)	48
Abbildung 6:	Zielprozess Produkterstellung	51
Abbildung 7:	Datenmodell Datenlieferungskonfiguration (delivery_configuration)	73
Abbildung 8:	Datenmodell Datenlieferungsprozess (delivery_process).....	74
Abbildung 9:	Datenmodell Listen (list).....	75
Abbildung 10:	Datenmodell Benutzer:innen (account)	76
Abbildung 11:	Gesamtdatenmodell WaBo-DMZ	77
Abbildung 12:	Kerndatenmodell FIS Wasser-Boden Intranet.....	79
Abbildung 13:	Datenmodell Stammdaten (master_data)	81
Abbildung 14:	Datenmodell Parameter (measurand)	82
Abbildung 15:	Datenmodell Messungen (measurement)	83
Abbildung 16:	Datenmodell Metadaten (metadata)	84
Abbildung 17:	Datenmodell Datenlieferungskonfiguration (delivery_configuration)	85
Abbildung 18:	Datenmodell Listen (list).....	86
Abbildung 19:	Datenmodell Benutzer:innen (account)	87
Abbildung 20:	Gesamtdatenmodell WaBo-Intranet	88
Abbildung 21:	Nachrichtenmodell Messdatenübermittlung Fließgewässer ...	92
Abbildung 22:	Nachrichtenmodell Messdatenübermittlung Seen	95
Abbildung 23:	Nachrichtenmodell Messdatenübermittlung Grundwasser.....	96
Abbildung 24:	Gemeinsames Nachrichtenmodell Wasser	97
Abbildung 25:	Nachrichtenmodell Messdatenübermittlung Boden, Hauptmodell	98
Abbildung 26:	Nachrichtenmodell Messdatenübermittlung Boden, Analyse und Methodencode	99

Tabellenverzeichnis

Tabelle 1:	Rollen von Benutzer:innen WaBo-Intranet	61
Tabelle 2:	Benutzer:innen-Rechte WaBo-Intranet.....	61
Tabelle 3:	Rollen von Benutzer:innen WaBo-DMZ.....	65
Tabelle 4:	Rollen von Organisationen	65
Tabelle 5:	Benutzer:innen-Rechte WaBo-DMZ	65

Abkürzungsverzeichnis

Abkürzung	Erläuterung
ARIT	Architekturrichtlinie für die IT des Bundes
ASCII	American Standard Code for Information Interchange
bBIS	bundesweites Bodeninformationssystem
BfG	Bundesanstalt für Gewässerkunde
BMU	Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz
BMWK	Bundesministerium für Wirtschaft und Klimaschutz
BPMN2	Business Process Model and Notation Version 2.0
BUBE	Betriebliche Umweltdatenberichterstattung
CSS	Cascading Style Sheets: Stylesheet-Sprache für elektronische Dokumente
DB	Datenbank
DESTATIS	Statistisches Bundesamt, Wiesbaden
DMZ	Demilitarisierte Zone: Subnetz zwischen dem öffentlichen Internet und internen lokalen Netzwerken einer Organisation
eBIS	elektronisches Bodeninformationssystem
EIONET	European Environment Information and Observation Network
ENDA	ENDA GmbH & Co. KG, Berlin
EQR	Ecological Quality Ratios
EUA	Europäische Umweltagentur
FG	Fachgebiet
FIS Wasser-Boden	Fachinformationssystem Wasser & Boden
ggü.	gegenüber
HELCOM	Baltic Marine Environment Protection Commission (Helsinki Commission)
HTTPS	Hypertext Transfer Protocol Secure
ICES	International Council for the Exploration of the Sea
i.d.R.	in der Regel
i.S.v.	Im Sinne von
IT	Information Technology (Informations-Technologie)
IT-NetzG	Gesetz über die Verbindung der informationstechnischen Netze des Bundes und der Länder
JSON	JavaScript Object Notation
KOM	Europäische Kommission

Abkürzung	Erläuterung
LANUV	Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen
LAWA	Bund/Länder-Arbeitsgemeinschaft Wasser
LUNG MV	Landesamt für Umwelt, Naturschutz und Geologie Mecklenburg-Vorpommern
LV	Liefervorgang
MUDAB	Meeresumweltdatenbank
NdB-VN	Netze des Bundes-Verbindungsnetz
NRW	Nordrhein-Westfalen
OGewV	Verordnung zum Schutz der Oberflächengewässer (Oberflächengewässerverordnung)
OSPAR	Convention for the Protection of the Marine Environment of the North-East Atlantic (Oslo Paris Convention)
OZG	Onlinezugangsgesetz
QS	Qualitätssicherung
QUADAWA	Qualitätsdatenbank Wasser
REST	Representational State Transfer (dt. „Repräsentativer Zustandstransfer“)
RZ	Rechenzentrum
s.o.	siehe oben
UBA	Umweltbundesamt Dessau-Roßlau
v.a.	vor allem
VwV	Verwaltungsvereinbarung
WFD	Water Framework Directive (deutsch: Wasserrahmenrichtlinie, WRRL)
WISE	Water Information System for Europe
WRRL	EU Wasserrahmenrichtlinie 2000/60/EG
XML	Extensible Markup Language
XÖV	XML in der öffentlichen Verwaltung

Zusammenfassung

Einführung und Hintergrund

Das Forschungsprojekt „Anpassung der Berichtsprozesse für Wasser und Boden – Entwicklung einer übergreifenden und integrierten Infrastruktur für Zustandsdaten im Boden- und Gewässerschutz“ (FKZ 3720122020) wurde vom Umweltbundesamt (UBA) initiiert, um die Berichtsprozesse für die Umweltmedien Fließgewässer, Seen, Grundwasser und Boden zu harmonisieren. Die bisherigen Prozesse waren durch heterogene Datenformate, unterschiedliche Datenstrukturen und isolierte Werkzeuge geprägt, was die Effizienz der Datenverarbeitung, die Vergleichbarkeit der Ergebnisse und die Erfüllung von Berichtspflichten gegenüber der Europäischen Kommission (KOM) und anderen Organisationen erschwerte. Qualitätsgesicherte Umweltdaten sind wichtig für Politikberatung, Forschung, die Ableitung von Schutzmaßnahmen und die Information der Öffentlichkeit, die zu den Aufgaben des UBA gehören.

Ziele des Forschungsvorhabens

Das Projekt verfolgte mehrere übergeordnete Ziele:

- ▶ **Vereinheitlichung:** Schaffung eines einheitlichen Fachinformationssystems (FIS Wasser-Boden), das die Datenstrukturen und Prozesse für Fließgewässer, Seen, Grundwasser und Boden konsolidiert, um Synergieeffekte zu nutzen und Redundanzen zu minimieren.
- ▶ **Interoperabilität:** Förderung der Zusammenarbeit mit anderen Systemen durch moderne Schnittstellen (z. B. REST-APIs) und standardisierte Datenformate wie .xlsx, CSV und JSON.
- ▶ **OpenData:** Bereitstellung qualitätsgesicherter Daten für die Öffentlichkeit und andere Institutionen über ein Data Warehouse und ein benutzerfreundliches Abfragewerkzeug.
- ▶ **Digitalisierung:** Unterstützung der digitalen Transformation durch automatisierte Datenverarbeitung, effiziente Datenabfragen und die Reduzierung von Netzwerkanforderungen für Remote Work.
- ▶ **Flexibilität:** Entwicklung einer skalierbaren Architektur, die die Integration weiterer Umweltmedien, wie Meeresumweltdaten (MUDAB), ermöglicht, ohne die Komplexität signifikant zu erhöhen.
- ▶ **Berichterstellung:** Optimierung der Berichterstellung für internationale Verpflichtungen, wie die EU-Wasserrahmenrichtlinie (WISE-6), das Global Environment Monitoring System (GEMS/Water) und die Nitratrichtlinie.

Darüber hinaus sollte das Projekt den Anforderungen der Architekturrichtlinie IT-Bund, des Barrierefreiheitsgesetzes und der XÖV-Standards entsprechen, um eine zukunftssichere Lösung zu gewährleisten.

Gewählter Ansatz

Das Projekt war in fünf Arbeitspakete gegliedert, die konzeptionelle, technische und forschungsbezogene Aspekte abdeckten:

- ▶ **Arbeitspaket 1 – Erarbeitung eines Fachkonzeptes:** Eine umfassende Analyse der bestehenden Datensätze, technischen Rahmenbedingungen und fachlichen Anforderungen wurde durchgeführt, um ein Konzept für das FIS Wasser-Boden zu entwickeln. Dieses Konzept zielte auf eine medienbruchfreie Verarbeitung von der Datenerfassung bis zur Produktgenerierung ab.

- ▶ **Arbeitspaket 2 – Umsetzung des Konzeptes:** Die technische Entwicklung umfasste die Erstellung eines Anwendungskerns, die Implementierung von Rollen- und Rechtekonzepten, die Entwicklung von Frontend- und Backend-Komponenten, die Integration in die UBA-IT-Umgebung, die Einrichtung von Datenbanken, die Migration von Altdaten sowie die Durchführung von Funktions- und Penetrationstests. Die Arbeiten wurden in enger Abstimmung mit den fachlichen Anforderungen der Umweltmedien durchgeführt.
- ▶ **Arbeitspaket 3 – Prüfung der Einbindung weiterer Umweltmedien:** Die Machbarkeit der Integration von Meeresumweltdaten (MUDAB) wurde untersucht, einschließlich einer Abschätzung der Aufwände und der notwendigen Anpassungen.
- ▶ **Arbeitspaket 4 – Integrierte Forschungsaspekte:** Forschungsfragen, wie die Nutzung externer Webdienste und die Entwicklung eines medienübergreifenden Ansatzes für die Datenrecherche, wurden im Kontext der Systementwicklung bearbeitet.
- ▶ **Arbeitspaket 5 – Abschlussdokumentation:** Ein umfassender Abschlussbericht wurde erstellt, der die Entwicklung, Implementierung und Ergebnisse des Projekts detailliert dokumentiert.

Die Umsetzung erfolgte nach agilen Methoden (Scrum), um Flexibilität und iterative Verbesserungen zu fördern. Regelmäßige Abstimmungen mit den Stakeholdern und die Einhaltung moderner IT-Standards stellten sicher, dass die Lösung den aktuellen und zukünftigen Anforderungen gerecht wird.

Zentrale Forschungsergebnisse

Einheitliche Datenstruktur und Prozesse

Das FIS Wasser-Boden vereinheitlicht die Erfassung, Verarbeitung und Analyse von Zustandsdaten für die genannten Umweltmedien. Durch die Harmonisierung der Datenmodelle und Prozesse wurde eine umfassende Schnittmenge geschaffen, die spezifische Anforderungen der einzelnen Medien berücksichtigt. Dies führte zu folgenden Ergebnissen:

- ▶ **Synergieeffekte:** Die gemeinsame Nutzung von Werkzeugen wie dem Abfragewerkzeug, dem Data Warehouse und Importmodulen reduzierte Entwicklungs- und Wartungsaufwände erheblich. Beispielsweise wurde die Importroutine für die Wasser-Kompartimente nur einmal entwickelt und mit spezifischen Erweiterungen für Boden angepasst.
- ▶ **Interoperabilität:** Moderne REST-Schnittstellen ermöglichen die einfache Integration mit anderen Systemen, z. B. den Datenbanken der Bundesländer. Unterstützte Formate wie .xlsx, CSV und JSON erleichtern den Datenimport und -export.
- ▶ **OpenData:** Qualitätsgesicherte Daten werden automatisiert im Data Warehouse abgelegt und über ein intuitives Abfragewerkzeug zugänglich gemacht, das auch automatisierte Abfragen und Downloads ermöglicht.

Technische Umsetzung

Das FIS Wasser-Boden wurde als Webanwendung entwickelt, die vollständig in die IT-Umgebung des UBA integriert ist. Es umfasst Module für:

- ▶ **Datenanzeige und -bearbeitung:** Benutzerfreundliche Oberflächen zur Verwaltung von Stammdaten, Messwerten und Metadaten.

- ▶ **Datenaggregation:** Automatische Berechnungen und Visualisierungen, z. B. Kurvendigramme, Histogramme und Boxplots.
- ▶ **Berichterstellung:** Exportfunktionen für standardisierte Berichte in XML- und CSV-Formaten.
- ▶ **Qualitätssicherung:** Automatische Prüfungen bei Datenimporten, z. B. auf Pflichtfelder, Datentypen und Plausibilitätsgrenzen, gewährleisten die Datenintegrität.
- ▶ **Erweiterbarkeit:** Das FIS Wasser-Boden kann zukünftig die derzeit bekannten Arten biologischer Parameter (Taxonomie und Abundanz, Makrophytenausdehnung, Chla- und Phaeophytindaten, Ecological Quality Ratios (EQRs) und Zustandsklassen des chemischen Zustands) aufnehmen und verarbeiten.

Die Datenmodelle wurden so gestaltet, dass sie sowohl die gemeinsamen als auch die spezifischen Anforderungen der Umweltmedien abdecken. Ein differenziertes Rollen- und Rechtekonzept sorgt für eine sichere und benutzergruppenspezifische Nutzung.

Digitalisierung und Flexibilität

Das System trägt wesentlich zur Digitalisierung bei:

- ▶ **Evidenzbasierte Entscheidungen:** Qualitätsgesicherte Daten unterstützen die Ableitung von Maßnahmen und die Beurteilung des Umweltzustands.
- ▶ **Öffentliche Anfragen:** Schnelle und kompetente Antworten auf Anfragen durch effiziente Datenabfragen.
- ▶ **Remote Work:** Reduzierte Netzwerkanforderungen im Vergleich zu traditionellen Systemen wie geteilten Laufwerken.
- ▶ **Automatisierung:** KI-basierte Ansätze, z. B. für die Konvertierung von Bodendaten, ermöglichen eine automatisierte Verarbeitung und sparen Ressourcen.

Die flexible Architektur des FIS unterstützt die Integration neuer Umweltmedien. Die Untersuchung der MUDAB-Integration offenbarte jedoch Herausforderungen, wie uneinheitliche Datenformate, unklare Dokumentationen und komplexe Zeichenketten in Parameterkennungen. Diese erfordern zusätzliche Datenforensik und Abstimmung mit der Bundesanstalt für Gewässerkunde (BfG), um eine erfolgreiche Integration zu gewährleisten.

Berichterstellung

Das FIS unterstützt internationale Berichtspflichten durch automatisierte Exportfunktionen, z. B.:

- ▶ **WISE-6:** Erzeugung von XML-Dateien für die Berichterstellung an die Europäische Umweltagentur (EUA).
- ▶ **GEMS/Water:** Generierung von CSV-Dateien für das Umweltprogramm der Vereinten Nationen. Die Konfiguration von Berichtsparametern erfolgt über die Weboberfläche, und QS-Prüfungen stellen sicher, dass die exportierten Daten den Anforderungen entsprechen.
- ▶ **Nitratrichtlinie:** Erstellung von .xlsx-Dateien mit für die Nitratberichterstattung spezifischer Struktur.

Herausforderungen und Limitationen

Die Integration von Meeresumweltdaten (MUDAB) erweist sich als komplex, da die vorhandenen Datenstrukturen und Dokumentationen teilweise inkonsistent sind. Beispielsweise fehlen klare Referenzen zwischen Messwerten und Parametern, und die Interpretation von Zeichenketten erfordert umfangreiche analytische Arbeit. Die Webpräsenz der MUDAB ist derzeit nicht nutzerfreundlich, was die Datenbereitstellung erschwert. Ungeachtet dessen wird eine Empfehlung zur Integration der MUDAB in das FIS Wasser-Boden ausgesprochen.

Ausblick

Das FIS Wasser-Boden bietet eine geeignete Grundlage für die zukünftige Umweltüberwachung und -berichterstattung. Die erfolgreiche Vereinheitlichung der Prozesse und Datenstrukturen zeigt, wie moderne IT-Lösungen die Effizienz und Transparenz in der Umweltverwaltung steigern können. Zukünftige Entwicklungen könnten den verstärkten Einsatz von KI, die Integration weiterer Umweltmedien und die Verbesserung der Webdienste umfassen. Die Integration der MUDAB bleibt eine Priorität, erfordert jedoch gezielte Anpassungen. Das Projekt unterstreicht die Bedeutung einer strategischen IT-Konsolidierung für die Erreichung umweltpolitischer Ziele.

Summary

Introduction and Background

The research project “Adaptation of Reporting Processes for Water and Soil – Development of a Comprehensive and Integrated Infrastructure for Status Data in Soil and Water Protection” (FKZ 3720122020) was commissioned by the German Environment Agency (UBA) to harmonise the reporting processes for the environmental media of rivers, lakes, groundwater, and soil. The previous processes were characterised by heterogeneous data formats, disparate data structures, and isolated tools, which hindered the efficiency of data processing, the comparability of results, and the fulfilment of reporting obligations to the European Commission (EC) and other organisations. Quality-assured environmental data are important for policy advice, research, the derivation of protective measures, and public information, which are responsibilities of the UBA.

Objectives of the Research Report

The project pursued several overarching objectives:

- ▶ **Standardisation:** Development of a unified Fachinformationssystem (FIS Water-Soil) to consolidate data structures and processes for rivers, lakes, groundwater, and soil, leveraging synergies and minimising redundancies.
- ▶ **Interoperability:** Enhancing collaboration with other systems through modern interfaces (e.g., REST APIs) and standardised data formats such as .xlsx, CSV, and JSON.
- ▶ **OpenData:** Providing quality-assured data to the public and other institutions via a data warehouse and a user-friendly query tool.
- ▶ **Digitalisation:** Supporting digital transformation through automated data processing, efficient data queries, and reduced network requirements for remote work.
- ▶ **Flexibility:** Creating a scalable architecture that allows the integration of additional environmental media, such as marine environmental data (MUDAB), without significantly increasing complexity.
- ▶ **Reporting:** Streamlining reporting for international obligations, such as the EU Water Framework Directive (WISE-6), the Global Environment Monitoring System (GEMS/Water) and the Nitrate Directive.

Additionally, the project aimed to comply with the Federal IT Architecture Guideline, the Accessibility Act, and XÖV standards to ensure a future-proof solution.

Chosen Approach

The project was structured into five work packages covering conceptual, technical, and research-related aspects:

- ▶ **Work Package 1 – Development of a Conceptual Framework:** A comprehensive analysis of existing datasets, technical constraints, and specific requirements was conducted to develop a concept for the FIS Water-Soil, aiming for seamless data processing from collection to product generation.
- ▶ **Work Package 2 – Implementation of the Concept:** Technical development included creating an application core, implementing role and rights concepts, developing frontend and backend components, integrating the system into the UBA’s IT environment, setting up

databases, migrating legacy data, and conducting functional and penetration tests. The work was closely aligned with the specific requirements of the environmental media.

- ▶ **Work Package 3 – Assessment of Integrating Additional Environmental Media:** The feasibility of integrating marine environmental data (MUDAB) was examined, including an estimation of efforts and necessary adaptations.
- ▶ **Work Package 4 – Integrated Research Aspects:** Research questions, such as the use of external web services and the development of a cross-media approach for data retrieval, were addressed in the context of system development.
- ▶ **Work Package 5 – Final Documentation:** A comprehensive final report was compiled, detailing the project's development, implementation, and outcomes.

The implementation followed agile methods (Scrum) to promote flexibility and iterative improvements. Regular coordination with stakeholders and adherence to modern IT standards ensured the solution met current and future requirements.

Key Research Findings

Unified Data Structure and Processes

The FIS Water-Soil standardises the collection, processing, and analysis of status data for the mentioned environmental media. By harmonising data models (WaBo-Intranet and WaBo-DMZ) and processes, a maximum common ground was established without neglecting the specific needs of individual media. This led to the following outcomes:

- ▶ **Synergies:** Shared use of tools such as the query tool, data warehouse, and import modules significantly reduced development and maintenance costs. For instance, the import routine for water-related media was developed once and adapted for soil with specific extensions.
- ▶ **Interoperability:** Modern REST interfaces enable seamless integration with other systems, such as state-level databases. Supported formats like .xlsx, CSV, and JSON facilitate data import and export.
- ▶ **OpenData:** Quality-assured data are automatically stored in the data warehouse and made accessible via an intuitive query tool that supports automated queries and downloads.

Technical Implementation

The FIS Water-Soil was developed as a web application fully integrated into the UBA's IT environment. It includes modules for:

- ▶ **Data Display and Editing:** User-friendly interfaces for managing master data, measurement values, and metadata.
- ▶ **Data Aggregation:** Automated calculations and visualisations, e.g., line charts, histograms, and boxplots.
- ▶ **Reporting:** Export functions for standardised reports in XML and CSV formats.
- ▶ **Quality Assurance:** Automated checks during data imports, e.g., for mandatory fields, data types, and plausibility limits, ensure data integrity.
- ▶ **Extensibility:** The FIS Water-Soil interface will be able to record and process the currently known types of biological parameters (taxonomy and abundance, macrophyte coverage,

chlorophyll-a and phaeophytin data, Ecological Quality Ratios (EQRs) and status classes of chemical condition) in the future.

The data models were designed to address both common and specific requirements of the environmental media. A differentiated role and rights concept ensures secure and user-group-specific access.

Digitalisation and Flexibility

The system significantly contributes to digitalisation by:

- ▶ **Evidence-Based Decision-Making:** Quality-assured data support the derivation of measures and the assessment of environmental conditions.
- ▶ **Public Inquiries:** Rapid and competent responses to inquiries through efficient data queries.
- ▶ **Remote Work:** Reduced network requirements compared to traditional systems like shared drives.
- ▶ **Automation:** AI-based approaches, e.g., for converting soil data, enable automated processing and save resources.

The flexible architecture of the FIS allows the integration of new environmental media. However, the assessment of MUDAB integration revealed challenges, such as inconsistent data formats, unclear documentation, and complex parameter identifiers. These require additional data forensics and coordination with the Federal Institute for Hydrology (BfG) to ensure successful integration.

Reporting

The FIS supports international reporting obligations through automated export functions, e.g.:

- ▶ **WISE-6:** Generation of XML files for reporting to the European Environment Agency (EEA).
- ▶ **GEMS/Water:** Creation of CSV files for the United Nations Environment Programme.
- ▶ **Nitrate Directive:** Creation of .xlsx Files with specific structure required by the nitrate reporting guidelines.

Report parameters are configured via the web interface, and quality assurance checks ensure compliance with requirements.

Challenges and Limitations

Integrating marine environmental data (MUDAB) proved complex due to inconsistent data structures and documentation. For example, there are missing references between measurement values and parameters, and interpreting concatenated strings requires extensive analytical work. The current MUDAB web presence is not user-friendly, complicating data provision. These challenges necessitate additional resources and coordination with the BfG.

Outlook

The FIS Water-Soil provides a robust foundation for future environmental monitoring and reporting. The successful standardisation of processes and data structures demonstrates how modern IT solutions can enhance efficiency and transparency in environmental administration. Future developments could include increased use of AI, integration of additional environmental media, and improvements to web services. MUDAB integration remains a priority but requires

targeted adaptations. The project underscores the importance of strategic IT consolidation for achieving environmental policy objectives.

1 Einleitung

Qualitätsgesicherte Umweltdaten bilden die Basis für die Erfüllung von Berichtspflichten gegenüber der Europäischen Kommission (KOM) und anderen Organisationen. Sie sind auch notwendig für die Ableitung und Bewertung von Maßnahmen zum Schutz der Umwelt. Für die Forschung, Politikberatung und Information der Öffentlichkeit, die im Umweltbereich originäre Aufgaben des Umweltbundesamts (UBA) sind, sind sie ebenfalls unerlässlich.

Die Daten der Umweltmedien Fließgewässer, Seen, Grundwasser und Boden werden im UBA erfasst. Die bislang dafür genutzten Prozesse und in den Prozessen verwendeten Formate waren ebenso unterschiedlich wie die internen Strukturen der Datenhaltungen, die Zugänge zu den Daten und die eingesetzten Werkzeuge für die Berichterstellung.

Wenngleich sich aufgrund der unterschiedlichen Natur der genannten Umweltmedien nicht alle Strukturen vereinheitlichen lassen, so besteht doch eine große Überschneidung. Dies gilt auch für die Prozesse, die sich auf die Prüfung und den Import der Eingangsdaten, die Nutzung durch den/die Fachanwender:in, den Export von Berichtsdaten und die Bereitstellung von Produkten erstrecken.

Im Rahmen dieses Forschungsvorhabens wurde ein Konzept zur Vereinheitlichung der Datenstrukturen und der Nutzungsprozesse der o.g. Umweltkompartimente entwickelt und anschließend in einer an den spezifischen Erfordernissen der UBA IT orientierten Architektur als Webanwendung realisiert.

Auch die Erweiterbarkeit um neue Fachverfahren und Umweltmedien wird dadurch vereinfacht, ohne die Nachteile einer generischen Anwendung — vor allem unspezifische Begrifflichkeiten, aber auch umständliche Bedienung — in Kauf nehmen zu müssen.

Der Zweck der Vereinheitlichung von Daten und Prozessen war nicht primär ein höheres Maß an Ordnung und einheitlicher Verarbeitungsweisen, sondern deren Folgeergebnisse: Die leichtere, ggf. temporäre Übertragung von Aufgaben zwischen Organisationsbereichen, Ersparnis durch die Entwicklung gemeinsam genutzter Werkzeuge anstelle von je Umweltkompartiment entwickelter Werkzeuge, Ersparnis durch reduzierte Pflege- und Hostingaufwände, vor allem aber ein einheitliches Erscheinungsbild und ein einheitlicher Zugang zu den Produkten.

Die Herausforderung bestand darin, die Prozesse und Datenmodelle mit einer umfassenden Gemeinsamkeit zu vereinigen und die individuellen Anforderungen abzubilden, ohne dass die Komplexität von Prozessen, Datenmodell und Benutzerführung wesentlich zunimmt.

2 Mit der Leistungsbeschreibung verfolgte Ziele und umzusetzende Aufgaben

Das Projekt 25 105 / 552, FKZ: 3720 12 202 0 (Ressortforschungsplan-Vorhaben) „Anpassung der Berichtsprozesse für Wasser und Boden - Entwicklung einer übergreifenden und integrierten Infrastruktur für Zustandsdaten im Boden- und Gewässerschutz“ zielt auf die Vereinigung der vier o.g. Fachverfahren, konkret die im UBA Abteilung II angesiedelten Fachverfahren für die Umweltmedien Fließgewässer, Seen, Grundwasser und Boden (Bodenzustand und Bodenmonitoring).

Dabei sollen sowohl die in den Fachverfahren realisierten Prozesse als auch deren elektronische Unterstützung so weit möglich und sinnvoll vereinheitlicht und in jedem Fall fachlich und technisch konsolidiert werden.

2.1 Arbeitspaket 1 - Erarbeitung eines Fachkonzeptes

Aufgabe gem. Leistungsbeschreibung: „Ziel des Arbeitspakets ist die konzeptionelle Erarbeitung des geplanten Fachinformationssystems (Fachkonzept). Nach einer umfassenden Analyse der zu integrierenden Datensätze, den technischen Randbedingungen und den individuellen fachlichen Erfordernissen, soll ein Fachkonzept zur Umsetzung des geplanten Fachinformationssystems erarbeitet werden. Das Konzept soll als Ziel die medienbruchfreie Verarbeitung und digitale Durchgängigkeit von der Datenbereitstellung über die Verarbeitung bis zur Erstellung eines Produkts beinhalten.“

2.2 Arbeitspaket 2 - Umsetzung des Konzeptes

Aufgabe gem. Leistungsbeschreibung: „Das Arbeitspaket 2 umfasst die vollständige Umsetzung des neuen Fachinformationssystems in den Wirkbetrieb und einen abschließenden Funktions- und Penetrationstest. Die beiden im folgenden beschriebenen Unterarbeitspakete sollen in enger Verzahnung abgearbeitet werden.“

Das Arbeitspaket untergliedert sich organisatorisch grob in folgende Teilarbeitspakete:

- ▶ Arbeitspaket 2.1 - Technische Umsetzung des Konzeptes und
- ▶ Arbeitspaket 2.2 - Integration der individuellen fachlichen Anforderungen der einzelnen Umweltmedien

und besteht technisch aus folgenden Arbeitspaketen:

- ▶ Erstellung Anwendungskern
- ▶ Implementation Rollen- und Rechtekonzept und API für Frontend
- ▶ Frontendkern und essenzielle Navigation
- ▶ Frontend Nutzer-, Gruppenverwaltung
- ▶ Implementation Datenbanken
- ▶ Deployment, Tests und Continuous Integration errichten
- ▶ Grafisches Anwendungsdesign
- ▶ Fachspezifische Frontend-Entwicklung und -Finish

- Implementation Data-Warehouse
- Altdatenmigration

2.3 Arbeitspaket 3 - Prüfung der Einbindung weiterer Umweltmedien

Aufgabe gem. Leistungsbeschreibung: „Neben den genannten Daten zu Fließgewässern, Seen, Grundwasser und Boden sollen perspektivisch auch die Meeresumweltdaten in das neue Fachinformationssystem integriert werden. Dafür soll im Rahmen dieses Arbeitspaketes geprüft werden, wie diese Daten und die bereits bestehenden Auswerteroutinen dafür in das neue Fachinformationssystem integriert werden können. Der Datenbankdump und die Beschreibung der entsprechenden Funktionalitäten werden dem Auftragnehmer im Laufe des Projektes zur Verfügung gestellt. Im Ergebnis der Prüfung sollen die Aufwände und Anpassungsbedarfe abgeschätzt und erörtert werden. Dieses AP kann am Ende des Projektes bearbeitet werden.“

2.4 Arbeitspaket 4 - Integrierte Forschungsaspekte und Umsetzung der Forschungsfragen

Aufgabe gem. Leistungsbeschreibung: „Neben der konzeptionellen und technischen Umsetzung des neuen Fachinformationssystems sollen im Projekt auch Forschungsfragen bearbeitet werden, die nachfolgend beschrieben werden. Diese sollen stets im Kontext der Umsetzung des gesamten Projektes mit behandelt und deren Implementierung im Abschlussbericht des Projektes in einem separaten Kapitel erörtert werden.“

Das Arbeitspaket 4 untergliedert sich in folgende Teilaufgaben:

- a) Nutzung fremder Webdienste
- b) Erarbeitung eines Ansatzes medienübergreifender Sicht und Recherche

2.5 Arbeitspaket 5 - Abschlussdokumentation

In diesem Arbeitspaket soll der Abschlussbericht inklusive der Dokumentation des Fachinformationssystems erstellt werden.

3 Übergeordnete Ziele

Neben der technischen Aufgabenstellung, das Datenmanagement für die Boden- und Gewässerzustandsdaten beim UBA neu aufzustellen und dabei fachlich und technisch zu konsolidieren, sind für eine zukunftsweisende Realisierung zahlreiche übergeordnete Ziele verfolgt worden. Dies waren neben den verbindlichen Vorgaben (z. B. Architekturrichtlinie IT Bund, Barrierefreiheitsgesetz ¹ und Einhaltung der XÖV-Standards) insbesondere folgende Themenfelder:

- ▶ Interoperabilität
- ▶ OpenData
- ▶ Schnittstellen
- ▶ Digitalisierung
- ▶ Synergieeffekte
- ▶ Agile Methoden
- ▶ Multikulturelle Teams
- ▶ Gender
- ▶ Vorteile der Zusammenarbeit mit verschiedenen Verwaltungen
- ▶ Ablösung veralteter Standards
- ▶ Flexibilität auch für Integration neuer Umweltfelder
- ▶ Integrative Betrachtung der Umwelt
- ▶ Unterstützung der Maßnahmenableitung
- ▶ Webdienste und deren Einordnung bezüglich der zukünftigen Bedeutung

Das vom UBA durchgeführte Vorhaben trägt der von der Bundesregierung durch zahlreiche Gesetze und Vorhaben derzeit forcierten Konsolidierung der IT-Landschaft Rechnung.

Dies zeigte sich in der Leistungsbeschreibung, die z. B. zur Erfüllung des IT-NetzG den Betrieb der gemeinsamen Anwendung zum Gewässer- und Bodenschutz im Rechenzentrum (RZ) des UBA vorsah, aber auch an vielen anderen Stellen, die die Vorgaben der grundlegend und aus unserer Sicht sehr erfolgreich überarbeiteten Architekturrichtlinie IT-Bund (Architekturrichtlinie) nicht nur aufgriff, sondern auch dem Wesen nach mit Leben erfüllte.

Dabei sind alle Ebenen der Architekturrichtlinie — die strategischen IT-architektonischen Ziele, die Rahmenarchitektur IT-Steuerung Bund und die Architekturgrundsätze, vor allem aber auch die übergreifenden und die konkreten Architekturvorgaben berücksichtigt und an keiner Stelle verletzt. Um diese Aussage treffen zu können, wurde die Architekturrichtlinie und ihr Anhang in allen Punkten mit der Leistungsbeschreibung und den internen Entwicklungsvorgehen in Konzeption und Entwicklung abgeglichen.

Im Folgenden sollen die übergreifenden Ziele bzgl. ihrer Realisierung und der dazu gewählten Ansätze weiter untersetzt werden.

¹Gesetz zur Umsetzung der Richtlinie (EU) 2019/882 des Europäischen Parlaments und des Rates über die Barrierefreiheitsanforderungen für Produkte und Dienstleistungen und zur Änderung anderer Gesetze vom 16.07.2021

Dabei weicht dieser Teil stilistisch vom nachfolgenden technischen Teil ab. Er ist jedoch für die Einordnung des Projekts in die Zielstellungen bzgl. der IT in Deutschland unerlässlich.

3.1 Interoperabilität

Im Bereich der Interoperabilität liegt einer der wesentlichen Fortschritte, die mit der Realisierung dieses Projekts erzielt wurde. Dies liegt nicht nur in der Integration der unterschiedlichen Wasser- und des Boden-Umweltkompartiments begründet. Es ist die Implementation des Abfragewerkzeugs, des Data Warehouse und der Einsatz von modernen, offenen REST-Schnittstellen, die jetzt und bei sich ändernden Aufgaben auch in Zukunft die Informationsbedürfnisse bzgl. aller in der Anwendung abgelegten Daten erfüllen werden.

REST-Schnittstellen

REST-Schnittstellen, auch bekannt als RESTful APIs (Representational State Transfer Application Programming Interfaces), sind ein Architekturstil für die Entwicklung von Webdiensten. REST nutzt die Standardprotokolle des Webs, insbesondere HTTP, um Daten zwischen Client und Server auszutauschen. REST-Schnittstellen lassen sich leicht spezifizieren. Aus ihrer Spezifikation wird üblicherweise automatisiert eine Online-Dokumentation erstellt. Sie erfreuen sich großer Beliebtheit, sind einfach und robust und ermöglichen die Überwindung von Medienbrüchen beim Austausch von Daten.

Auch die vereinheitlichte Importschnittstelle, die den Möglichkeiten der Datenlieferant:innen Rechnung trägt, steigert die Interoperabilität mit anderen Systemen, z. B. Systemen der Bundesländer.

Hier kommt dem Projekt zugute, dass die erhebliche Modernisierung der Architekturrichtlinie IT-Bund bzgl. der einzusetzenden technischen Mittel den Einsatz nicht nur moderner, sondern auch zukünftig tragender Technologien ermöglicht hat. Davon wurde Gebrauch gemacht.

3.2 OpenData

Für die Bereitstellung von Daten im Sinne von OpenData sind verschiedene Teile der Anwendung und ihr architektonisches Zusammenspiel von Bedeutung.

Das leicht und intuitiv bedienbare und trotzdem mächtige Abfragewerkzeug ermöglicht die automatisierte Ausführung von Abfragen und die ebenso automatisierte Ablage im für alle Mitarbeiter des UBA offenen Data Warehouse. Somit werden die mit dem qualitätsgesicherten Import in das FIS Wasser-Boden aufgenommenen Daten auch automatisch, in frei konfigurierbarem Turnus, in leicht verarbeitbarem Format bereitgestellt.

Diese Ergebnisse können per Download geladen werden und der Link auf das Ergebnis kann kopiert und mit dem jeweils geeigneten Kommunikationskanal verteilt werden.

Auch die nicht sofort ersichtlichen Voraussetzungen für die Realisierung leistungsfähiger Schnittstellen sind eingehalten worden. Der trotz der heterogenen Struktur der Umweltkompartimente homogene, klare und bzgl. der Datenklassen untereinander vollständig und mit sauberen Referenzen verdrahtete Modellkern bietet einerseits gute Performance bei der Datenextraktion und ermöglicht andererseits auch Dritten, die z. B. die Datenbank direkt nutzen wollen, sofort das für die Nutzung notwendige Verständnis über den inneren Aufbau von FIS Wasser-Boden.

3.3 Schnittstellen

Importseitig werden die Tabellenformate .xlsx und CSV unterstützt. Für den Export von z. B. Abfrageergebnissen kann auch das JSON-Format verwendet werden.

JSON-Format

JSON (JavaScript Object Notation) ist ein einfaches, aber mächtiges Format, das zur Übertragung von automatisiert verarbeitbaren Inhalten zwischen zwei Computern verwendet wird. Für den menschlichen Betrachter ist es leicht lesbar und auch leicht zu erstellen. Ein weiterer Vorteil ist, dass es mit Hilfe von Software sowohl leicht zu analysieren als auch zu erstellen ist. JSON besteht aus Schlüssel-Wert-Paaren und unterstützt grundlegende Datentypen wie Zeichenketten, Zahlen, Arrays und Objekte.

Darüber hinaus verfügt die Anwendung über moderne, klare und von beliebigen Anwendungen automatisiert nutzbare REST-Schnittstellen für jedwede Art von Information, die zwischen der Anwendung und der Benutzungsoberfläche (User Interface) ausgetauscht wird.

Die Benutzungsoberfläche ist eine eigene Anwendung, die ausschließlich für die Visualisierung im Webbrowser zuständig ist. Alle Teile dieser Anwendung kommunizieren über zahlreiche REST-Schnittstellen mit dem „Business Logic“ genannten Anwendungskern.

Es existieren folglich Schnittstellen für die Extraktion und die Änderung von Stammdaten, Messungen und Messwerten, für die Erstellung und Ausführung von Abfragen, für das Auslesen aller Inhalte des Data Warehouse und das Hinzufügen sowie die Erstellung von Ordnern und somit beliebigen Strukturen im Data Warehouse. Daneben existieren Schnittstellen für das Auslesen, Ändern und die Erstellung von Messnetzen und auch von Parametergruppen, die beide sowohl in Abfragen als auch in der Reporterstellung genutzt werden können.

Diese Schnittstellen gehorchen einem klaren und verständlichen Benennungsschema, sind in ihrem Aufbau einheitlich und so implementiert, dass ihre Nutzung möglichst einfach erfolgt. Sie entsprechen dem aktuellen Stand von Forschung und Technik und erfüllen die Anforderungen der Architekturrichtlinie.

3.4 Digitalisierung

Für die Untersuchung, in welchen Bereichen FIS Wasser-Boden einen Beitrag für die Digitalisierung leistet, bedurfte es zunächst einer Begriffsbestimmung und der Ableitung von Kriterien (s.u., Kap. 3.4.1). Dazu wurden die nicht-kommerziellen, also die nicht durch Werbung auf die ersten Plätze beförderten Ergebnisse einiger Suchmaschinen ausgewertet, um zu Kriterien zu gelangen, woran sich die Digitalisierung in Deutschland fest macht. Zu folgenden Digitalisierungsthemen leistet FIS Wasser-Boden einen aktiven oder passiven Beitrag:

- ▶ Administrativ-rechtliche Rahmenbedingungen
 - Hier hilft das FIS Wasser-Boden durch die Bereitstellung qualitätsgesicherter Daten bei der Ableitung von evidenzbasierten Handlungsempfehlungen und Maßnahmen. Auch die Beurteilung des Zustands von Umweltkompartimenten, die für die Identifikation von Handlungsfeldern wichtig ist, wird unterstützt.
- ▶ Gesellschaft
 - FIS Wasser-Boden ermöglicht es, bei Anfragen mit Datenhintergrund schnell und kompetent Auskunft zu geben.

► digitale Verwaltungsleistungen

- Hier ist das FIS Wasser-Boden derzeit indirekt beteiligt, jedoch aufgrund des Datenumfangs von ca. 10 Millionen Messwerten bei gleichzeitig klarer Strukturierung ein Datenlieferant für Projekte in der ersten Reihe, wie die Umweltinformationsplattform Data Cube, die unter <https://datacube.uba.de> abrufbar ist und die Geodateninfrastruktur des Umweltbundesamts (Metadatenkatalog des UBA), die unter <https://gis.uba.de> erreichbar ist.

► Remote Work

- FIS Wasser-Boden reduziert gegenüber der Nutzung von geteilten Laufwerken mit Tabellen- oder Access-Dateien die benötigte Netzwerkbandbreite erheblich und ist daher günstig in Bezug auf Arbeiten im Homeoffice

Zukünftig könnte das FIS Wasser-Boden auch künstliche Intelligenz, zumindest in erweitertem Sinn, einsetzen. Dies betrifft die Aufnahme von Bodendaten, für die bereits ein Training mit tausenden von Eingangsdatensätzen vorgenommen wurde. Die daraus gewonnenen Erkenntnisse werden eingesetzt, um die Konvertierung von lieferantenspezifischen Listeneinträgen zu in FIS Wasser-Boden bekannten Listeneinträgen automatisiert vorzunehmen. Das ermöglicht die automatisierte Verarbeitung von Eingangsdaten und somit überhaupt die Aufnahme eines großen Teils von Bodendaten und spart Zeit und Arbeitskraft.

3.4.1 Ableitung der Kriterien zur Digitalisierung

Für die Ableitung von Kriterien zur Beurteilung des Beitrags von FIS Wasser-Boden zur Digitalisierung Deutschland wurden die weit verbreiteten Internet-Suchmaschinen Google, Bing und Ecosia mit dem jeweils gleichen Suchbegriff „Digitalisierung Deutschland“ verwendet. Dabei wurden die kommerziell beworbenen Abfrageergebnisse aus der Antwortmenge entfernt und jeweils die ersten drei nicht-kommerziellen Ergebnisse näher betrachtet.

► Google (Google LLC — Suche, o. J.)

Die Google-Suche erbrachte Hyperlinks auf die nachfolgend aufgeführten Webseiten:

- Digitalisierungsindex (Bundesministerium für Wirtschaft und Klimaschutz, o. J. - a), insbesondere
*„Messung der Digitalisierung der Wirtschaft
 Der Digitalisierungsindex misst den Stand der Digitalisierung der Wirtschaft am Standort Deutschland mithilfe von 36 Indikatoren. Diese bilden jeweils verschiedene Dimensionen der Digitalisierung ab. Sie zeigen damit ein umfassendes Bild des Status quo der Digitalisierung von Unternehmen sowie des Umfelds, in dem diese aktiv sind. Aggregiert ergeben die Indikatoren einen Indexwert für die Digitalisierung der Wirtschaft in Deutschland insgesamt.“*

Der Digitalisierungsindex ist in mehrere Ebenen unterteilt und weist auf der Ebene der Kategorien sowohl unternehmensinterne als auch unternehmensexterne Kategorien auf. Da hier nicht primär die Wirtschaft profitieren soll, wurden nur die unternehmensexternen Indizes betrachtet. Diese sind:

- Technische Infrastruktur
- Administrativ-rechtliche Rahmenbedingungen
- Gesellschaft

- Verfügbare Arbeitsleistung²
- Innovationslandschaft
- Strategie für einen digitalen Aufbruch (Bundesregierung, o. J.)

„Die Digitalstrategie für Deutschland ist der ‚Wegweiser für den digitalen Aufbruch‘. Alle Ministerien und das Kanzleramt haben sie mit vereinten Kräften erarbeitet – koordiniert vom Bundesministerium für Digitales und Verkehr. Die Strategie soll Deutschland digital voranbringen und formuliert Ziele, an denen sich die Regierung messen lassen will.“

Als wichtigste Kriterien werden genannt:

- Versorgung mit Glasfaseranschlüssen
- digitale Verwaltungsleistungen
- Digitalisierung. Den digitalen Wandel gestalten (Bundesministerium für Wirtschaft und Klimaschutz, o. J. - b)

„Die Digitalisierung begegnet uns überall – und hat die Art, wie wir arbeiten und leben, stark verändert. Videokonferenzen sind in vielen Berufen längst üblich, Menschen bestellen Lebensmittel immer mehr online, Urlaub buchen sie zunehmend im Internet. Der technologische Wandel ist in vollem Gange. ...“

Als Eckpunkte werden im einleitenden Fließtext zuvorderst genannt:

- Videokonferenzen
- Bestellen von Lebensmitteln
- Buchen des Urlaubs
- 3D-Drucker
- Künstliche Intelligenz

► Bing (Microsoft Corporation — Bing-Suche, o. J.)

Die Bing-Suche erbrachte Hyperlinks auf die nachfolgend aufgeführten Webseiten:

- Strategie für einen digitalen Aufbruch (Bundesregierung, o. J.)
Diese Webseite wurde zuvor im Rahmen der Google-Suche beschrieben.
- Digitalisierung - wo steht Deutschland? (Statista, o. J.)

„Wir leben in Zeiten einer zunehmenden Digitalisierung von Gesellschaft, Wirtschaft und Verwaltung. Dies stellt uns alle vor Herausforderungen und bietet zugleich große Chancen. Die erfolgreiche Gestaltung des digitalen Wandels wird die Weichen für die Zukunft stellen. Die Bundesregierung identifiziert in ihrer Digitalstrategie fünf Handlungsfelder. Digitale Kompetenzen der Gesellschaft stärken, eine funktionierende Infrastruktur und Ausstattung zur Verfügung stellen, Innovation und digitale Transformation fördern, den Menschen in den Mittelpunkt des digitalen Wandels stellen und alle Verwaltungsleistungen digital anbieten.“

² Der Originalbegriff „Humankapital“ wurde durch den Begriff „verfügbare Arbeitsleistung“ ersetzt

95 Prozent der Bevölkerung sind online ...“

Als wichtigste Kriterien werden genannt:

- Zahl der Internetnutzer
- Nutzung digitaler Dienste

Statista ist ein kommerzieller Dienst. Im öffentlich einsehbaren Teil waren keine weiteren, stichhaltigen Kriterien ersichtlich.

- Digitalisierung in Deutschland: Wie ist der aktuelle Stand? (Aganina, o. J.)

„Wie digital ist Deutschland?

‚Deutsche Wirtschaft verpasst die Digitalisierung‘, ‚Wie Deutschland die Digitalisierung verschläft‘ oder ‚Deutsche Arbeitnehmer nicht für Digitalisierung gewappnet‘: Diese und ähnliche Schlagzeilen kann man seit Jahren in den Medien hören und lesen.

Ist die Lage tatsächlich so fatal?

...

Was ist mit ‚der Digitalisierung‘ überhaupt gemeint?

Es gibt nicht die Digitalisierung. Was unter Digitalisierung, Digitaler Wandel oder auch Industrie 4.0 verstanden wird, ist ein großer Themenkomplex, der aus vielen Aspekten besteht. Zum Beispiel gehören dazu der Breitbandausbau, E-Learning, E-Government, E-Commerce, E-Health, Remote Work und die Nutzung von Cloud-Anwendungen. Ebenso zählen unter anderem Big Data, Künstliche Intelligenz, Smart Home, Connected Cities, Fintech, Predictive Maintenance, Deeptech, autonom fahrende Autos und Virtual Reality dazu.“

Die Kriterien zur Bewertung der Digitalisierung in Deutschland werden explizit genannt. Diese sind:

- Breitbandausbau
- E-Learning
- E-Government
- E-Commerce
- E-Health
- Remote Work

- ▶ Ecosia, DuckDuckGo

Diese Internet-Suchmaschinen liefern dieselben Ergebnisse in derselben Reihenfolge wie Bing. Daher erbringen sie keine weiteren Kriterien zur Beurteilung des Beitrags von FIS Wasser-Boden zur Digitalisierung in Deutschland.

3.5 Synergieeffekte

Mit dem FIS Wasser-Boden werden Synergien vor allem im Bereich der gemeinsamen Nutzung von nur einmal zu erstellenden Werken bzw. nur einmal zu erbringenden Leistungen erzielt. Dies betrifft das Hosting (eines statt eines je Umweltkompartiment) aber auch die Entwicklung besonderer Anwendungsteile wie des Abfragewerkzeugs, des Data Warehouse aber auch der

funktional mächtigen Messnetz- und Parametergruppen-Werkzeuge, die von allen Umweltkompartimenten genutzt werden können.

Offensichtlich ist es auch bei kompartimentübergreifenden Fragestellungen einfacher, die benötigten Daten aus einer Anwendung und in einem Format bereit zu stellen, als diese einzeln aus unterschiedlichen Fachanwendungen zusammenzustellen.

Es ergeben sich weitere Synergieeffekte bei der Entwicklung von QS-Prüfungen, da diese in der Regel ebenfalls in allen Umweltkompartimenten eingesetzt werden können. Dies betrifft z. B. die Prüfung von Messwerten auf Einhaltung von Wertebereichen je Messgröße, die Methoden zur Prüfung der Existenz einer Angabe und auch zu Prüfungen von abhängigen Feldern. (Wenn A angegeben ist, muss auch B angegeben sein; Wenn C angegeben ist, darf D nicht angegeben sein etc..)

Auch die Importroutine ist, zwar mit speziellen Erweiterungen je Kompartiment, für die Wasser-Kompartimente nur einmal entwickelt. Weiterhin können Teile der Importfunktionalität auch für die Bodendaten verwendet werden.

Die Ausführungen zu den Synergieeffekten sind sicher nicht vollständig, überlappen sich jedoch auch mit anderen Teilen wie z. B. der Thematik der Digitalisierung.

Leicht kann der Eindruck entstehen, dass es aufgrund der Synergien sinnvoll wäre, viele weitere, vielleicht auch umweltfremde, aber digitalisierbare Themen, mit in eine Anwendung aufzunehmen. Wächst jedoch damit die Komplexität, vielleicht linear, vielleicht sogar schneller, so wirkt sich dies ungünstig auf die Wartbarkeit aus. Es kann sich bei komplexer werdendem User Interface auch negativ auf die Nutzbarkeit auswirken. Negative Auswirkungen auf die Performance sind in dem Zusammenhang zwar zu erwarten, jedoch nicht in signifikantem Ausmaß.

Gelingt es jedoch, die Komplexität gleich zu halten, was für die Einbindung weiterer Themen mit messstellengebundenen Messwerten z. B. problemlos möglich ist, so entfallen alle zuvor als potenziell ungünstig aufgeführten Effekte und alle bekannten Synergien werden voll genutzt.

3.6 Agile Methoden

Der Einsatz agiler Methoden gilt seit einigen Jahren als probates Mittel, um starre und langsam agierende oder reagierende Organisationsstrukturen und vor allem deren Prozesse umzuformen, um höhere Flexibilität und schnellere Entscheidungen und Handlungen sowie eine Steigerung der Innovativität zu erreichen (Diehl, 2023; Moonkamp, 2024; Hunt et. al., 2001).

Dabei werden iterative Vorgehensweisen (z. B. Scrum) eingesetzt, die eine kurzzyklische Schaffens- und Bewertungsphase umsetzen.

Dabei widerstreben jedoch agile Methoden regelmäßig mit den Anforderungen der allermeisten Organisationen an von vornherein festgelegte Kosten bzw. Aufwände und klar definierte Leistungen (IT-P GmbH, 2021; Codefirst, 2017; Krusche & Company GmbH, 2020). Weiterhin lassen sich agile Methoden leichter in ohnehin gegenüber neuen Ansätzen aufgeschlossenen Teams umsetzen.

In diesem Projekt wurde Scrum mit großem Erfolg in der Entwicklungsphase eingesetzt, wobei inzwischen bekannte Fehler vermieden wurden. Eine umfangreiche Darstellung möglicher Fehlanwendungen von Scrum gibt Maggiori:

„Wie das agile Konzept die Produktivität tötet

Die Grundprinzipien der agilen Softwareentwicklung sind lobenswert und viel besser als die alten Ideen. Daher bemühen sich viele Unternehmen, agil zu sein. Sie übernehmen dafür agile Konzepte,

also Schritt-für-Schritt-Leitlinien, die ein Team agil machen sollen. Das berühmteste ist Scrum. Viele Unternehmen glauben, sie würden agil, indem sie sich strikt an eine Reihe unflexibler Regeln halten. Das Gegenteil passiert. (Maggiori, 2023)“

Dazu wurden folgende Vorgehensweise eingesetzt:

- ▶ Gemeinsame Aufstellung der Arbeitspakete (in Form von Issues/Tickets) für die nächste Sprint-Phase, aber vermeiden von Aufwandsfestlegungen
- ▶ Aufteilung der Arbeiten auf die jeweils am besten geeigneten Mitarbeitenden durch das Team selbst
- ▶ Vor der Review-Phase ist eine get-together-wir-helfen-uns-alle-gegenseitig-Phase gelebt worden. In dieser wurden diejenigen Aufgaben, die die Bearbeitenden erwartet hätten, in der Sprint-Phase erledigen zu können in kleinen gemeinsamen Teams bearbeitet und gelöst. Dies führt zu aktiver Know-How-Übertragung und vermeidet demotivierende Arbeitssituationen
- ▶ Nachfolgend eine gemeinsame Review-Phase, in der die getrennten Entwicklungsarbeiten zusammengeführt (gemergt), die Arbeiten getestet und die Issues/Tickets für die erledigten Arbeiten geschlossen wurden

Wichtig ist in diesem Zusammenhang, dass agile Methoden durch eine Überladung mit Verwaltungs- und Vorgehensvorschriften (i.S.v. „Wie macht man Agilität richtig — 17 Schritte, die eingehalten werden müssen“) erstickt werden. Agilität entsteht nicht, in dem man in alte Verhaltensmuster zurückfällt und alles Mögliche reglementiert, überwacht und auf mehreren Ebenen steuert.

Agilität entsteht in einem Team, das gemeinsame Ziele verfolgt und dabei nicht durch falsch eingesetzte Bonussysteme gegeneinander arbeitet. Es sollte so wenig steuernde Eingriffe wie möglich geben, und jedes Teammitglied sollte mit dem optimalen Werkzeugsatz ausgestattet werden. Dabei sollten persönliche Vorlieben berücksichtigt werden, denn diese dienen in aller Regel nur dem Ziel, die Arbeit besser und/oder schneller erledigen zu können.

3.7 Multikulturelles Team

Das ausführende Unternehmen nutzt das Potential eines multikulturellen Entwicklerteams, um durch verschiedenste Herangehensweisen an die Aufgabenstellungen breitgefächerte Lösungsansätze zu verfolgen. Dies ermöglicht die Nutzung kreativer Ressourcen für das Erreichen eines hohen Projekterfolgs.

Um diese Erfolge mit einem multikulturellen Team zu erreichen, ist die Projektsprache Englisch. Dies spiegelt sich auch in den Bezeichnungen aller technischen Ergebnisse wider.

3.8 Vorteile in der Zusammenarbeit mit verschiedenen Verwaltungen

Die Einbindung unterschiedlicher Verwaltungseinheiten bringt, bedingt durch den unterschiedlichen Erfahrungshorizont, auch unterschiedliche Lösungswege hervor. Die Realisierung dieser Ansätze in der Anwendung, erhöht deren Funktionsumfang. Dabei werden Nutzungsbarrieren beseitigt und für die anderen Verwaltungseinheiten auch neue Möglichkeiten der Nutzung geschaffen.

In diesem Projekt betraf dies konkret die Methodik der Datenaggregation, des Zugriffs auf Daten über spezifische Zugänge aber auch die Einbeziehung moderner Dienste (hier: WFS³ für Grundwasserdaten des LUNG MV) zur Datenakquisition.

3.9 Ablösung veralteter Standards

Das Vorhaben FIS Wasser-Boden hat durch die Zusammenführung der jeweils besten Standards der einzelnen Umweltkompartimente alle beteiligten Kompartimente auf das neueste, technische aber auch prozess-organisatorische Niveau gehoben. Dies betrifft die eingesetzten technischen Komponenten (DV-Architektur, Entwicklungsplattform, Datenbank) aber eben auch die Prozesse.

Als Beispiel ist für dieses Projekt die Organisation der Datenakquisition zu nennen, die in der Vergangenheit versucht hat, auch sich selten oder nicht ändernde Stammdaten über komplexe, teure und letztendlich vom Datenlieferanten schwer zu bedienende Importschnittstellen aufzunehmen.

Solche Daten werden zweckmäßig von den Fachverantwortlichen über moderne Oberflächen eingepflegt und können über diese auch gepflegt werden, so es denn notwendig wird.

Weiterhin ist eine konsequente Ausrichtung an den Vorgaben des XÖV, die klare, semantisch eindeutige und immer auch bzgl. ihrer Bedeutung beschriebener Bezeichner ebenso fordert, wie dokumentierte Prozesse, in FIS Wasser-Boden erfolgt. Dies ermöglicht die mit dem XÖV angestrebten Ziele der Nachvollziehbarkeit und vor allem des einfachen, offenen und deterministischen Datenaustauschs zu erreichen.

Nicht zuletzt durfte in diesem Projekt aktuelle, leistungsfähige, lizenzkostenfreie und quelloffene Software eingesetzt werden. Damit tragen wir auch den gesetzlichen Anforderungen Rechnung (Bundesministerium des Innern und für Heimat, 2017).

3.10 Flexibilität auch für Integration neuer Umweltfelder

Generell können neue Umweltfelder immer auch in bestehende Lösungen integriert werden. Es ist möglich, neue Inhalte neben die alten zu stellen. Das ist jedoch nur eine Schein-Integration, denn sie erhöht die Komplexität aller Anwendungsteile mindestens linear mit dem strukturellen Zuwachs.

Das kürzlich abgelöste BUBE-System zur Berichterstattung der Emissionsdaten nach e-PRTR-Verordnung, 11. BImSchV (Emissionserklärungsverordnung), 13. Und 17. BImSchV hat den Ansatz gewählt, alle Datenmodelle getrennt nebeneinander zu stellen. Da Unternehmen häufig nach mehreren der o.g. Rechtsquellen berichtspflichtig sind, entsteht hoher Aufwand beim Abgleich der Daten untereinander. Ebenfalls hoch ist der Aufwand, wenn Daten extern bereit gestellt werden müssen, da eine Anwendung erstellt werden muss, die auf Anforderung die Daten zum Zweck der Bereitstellung zusammenführt. Die Zusammenführung von Daten aus verteilten Datenmodellen ist immer eine komplexe Aufgabe und die dafür verantwortlichen Anwendungsteile sind bei Änderungen an der Struktur pflegeintensiv.

FIS Wasser-Boden erreicht die Flexibilität dadurch, dass mit einmaligem analytischem Aufwand die gemeinsamen Teile identifiziert und nur einmal, in einem gemeinsamen Teil abgelegt werden. Damit wächst die Komplexität und damit der Pflegeaufwand weniger als linear und die bereits vorhandenen Werkzeuge können weiter genutzt werden.

³ Web Feature Service: Elektronischer Dienst für automatisierte Datenübertragung

Daneben vermeidet FIS Wasser-Boden überbordende Anforderungen an das Datenmodell, in dem unnötige oder durch vereinfachte Prozesse vermeidbare Komplexität unterbleibt. Das System erfüllt alle gestellten Anforderungen und vermeidet unnötige Strukturen. Dies hat zur Folge, dass der bestehende Modellkern übersichtlich ist und bleibt und dadurch neue Umweltdaten besonders leicht integrierbar sind.

3.11 Integrative Betrachtung der Umwelt

Das in diesem Vorhaben entwickelte FIS Wasser-Boden ist durch die Zusammenführung auch sehr unterschiedlicher Umweltkompartimente⁴ ein gutes Beispiel für die integrative Betrachtung der Umwelt. Bereits in der Konzeptionsphase konnte das Entwicklungsteam die Vorteile nutzen, die aus der Kenntnis der in den jeweiligen Kompartimenten untersuchten Stoffe und deren Konzentrationen erwachsen. Dabei konnte jeweils ein Interessenausgleich und eine gemeinsame Sicht auf die Fachlichkeit erreicht werden.

Es war eine der wichtigen Anforderungen an FIS Wasser-Boden, dass sich weitere Umweltkompartimente möglichst leicht hinzufügen lassen. Dies wurde mit den bereits aufgeführten technischen Konzepten realisiert und somit kann die integrative Betrachtung auch weiter ausgebaut werden.

3.12 Unterstützung der Maßnahmenableitung

Die Ableitung von Maßnahmen zum Schutz der Umwelt folgt sinnvollerweise einer Kosten-Nutzen-Analyse (Umweltbundesamt, 2013). Bezüglich der Kostenfrage ist es wichtig zu wissen, an welchen Stellen Maßnahmen erforderlich sind, um die geologischen, hydrologischen und wirtschaftlichen Rahmenbedingungen einschätzen zu können. Dies kann durch die mit FIS Wasser-Boden generell verorteten Messstellenbezüge unterstützt werden.

Bezüglich der Wirksamkeit bzw. der Größe des Nutzens sind vorrangig die Konzentrationswerte, oft auch in Kombination mit weiteren Werten wie pH-Wert oder durchschnittlicher Temperatur, entscheidend. Auch hier liefert FIS Wasser-Boden die Datengrundlage. Nachfolgend kann die zu verfolgende Strategie zum Erreichen des jeweiligen Schutzzieles abgeleitet werden.

3.13 Webdienste

Das Fachinformationssystem FIS Wasser-Boden setzt auf allen Ebenen modernste Webdienste (RESTful Services) ein und kann seine Daten (Beispiel: Grundwasserdaten des LUNG MV) über moderne, standardisierte Webdienste beziehen.

Aus organisatorischer Sicht ist die Übermittlung von Daten über Webdienste wie oben genannt, zukünftig ein vielversprechender Weg. Wie auch bei anderen Übertragungsprozessen mit Webdiensten ist gerade in der Interbehördenkommunikation die automatische Übertragung ohne zwischengelagerte, administrative und manuelle Prozesse ein entscheidender Effizienztreiber.

Die Vorteile, die sich aus dem Einsatz von Webdiensten ergeben, sind vielfältig. Für die konkrete Anwendung FIS Wasser-Boden sind diese Vorteile IT-architektonischer Natur. Die Webdienste trennen verschiedene Anwendungsbereiche vollständig, was nicht nur die Wartbarkeit der Anwendung vereinfacht, sondern es auch ermöglicht, einzelne Teile mit minimalem Aufwand zu ersetzen, sollte die Notwendigkeit eintreten.

⁴ Die Wasserkompartimente (Grundwasser, Fließgewässer, Seen) auf der einen Seite und Boden auf der anderen Seite sind aus Umweltsicht sehr unterschiedlich

Für andere UBA-interne Fachinformationssystem ist die Unterstützung von Web-Diensten durch FIS Wasser-Boden von Vorteil, weil diese sich ohne die eingangs erwähnten, zwischengelagerten, administrativen und manuellen oder nur teilautomatisierte Prozesse der Daten des FIS Wasser-Boden bedienen können.

Dieser Vorgang, die benötigten Daten zu laden, kann vollautomatisiert und periodisch geschehen. Dabei können die Daten für die Veröffentlichung automatisiert aufbereitet oder um vertrauliche Daten reduziert werden. Es ist auch möglich, Webdienste „on demand“ auszuführen, also per Mausklick Daten abzurufen, wenn sie benötigt werden. Auch in diesem Fall werden Medienbrüche bei der Datenübertragung vermieden.

Das FIS Wasser & Boden ist zwar nicht öffentlich zugänglich, es führt jedoch alle Daten der Umweltkompartimente Grundwasser, Seen, Fließgewässer und Boden auf einer Plattform zusammen und ist auf die Erweiterbarkeit um weitere Kompartimente wie z.B. Meeresumweltdaten oder Pflanzenschutzmittel ausgelegt.

4 Anforderungen an Konzept und Umsetzung des FIS Wasser-Boden

Das hier vorliegende Konzept beschreibt in verschiedenen Ebenen die Anforderungen und Strukturen, die für die Entwicklung und Implementation des FIS Wasser-Boden eingehalten und umgesetzt werden mussten, um konform mit den organisatorischen Randbedingungen ein reibungsarm funktionierendes, leistungsfähiges, robustes und erweiterbares Fachinformationssystem im UBA zu implementieren.

Die wesentlichen Ebenen sind:

- ▶ IT-Umgebung des UBA
- ▶ Nutzungsprozesse
 - Der einzelnen Fachsysteme und
 - Ableitung eines vereinigenden Nutzungsprozesses FIS Wasser-Boden
- ▶ Datenmodelle
 - Der einzelnen Fachsysteme und
 - Ableitungen eines vereinigten Datenmodells für das FIS Wasser-Boden

Für die Konzepterstellung der gemeinsamen Fachanwendung sind alle Anforderungen bzgl. der Importschnittstellen, der Qualitätssicherung, der Möglichkeit der Datenrecherche (Abfragen), des Datenexports für beispielsweise verschiedene Berichtszwecke und der internen Dokumentationsanforderungen (Metadaten), die bekannt sind oder sich aus dem Kontext der Analyse ergeben, berücksichtigt worden.

Die technische Umsetzung und Integration in die IT-Umgebung des UBA musste die organisatorisch-technischen Anforderungen des UBA Rechenzentrums berücksichtigen und dabei eine Architektur und eine technische Basis entwerfen, die es möglichst effizient ermöglicht, eine mit den Anforderungen der Fachanwender:innen konforme IT-Lösung zu entwickeln, im UBA zu installieren, zu betreiben und auch zu pflegen.

4.1 Technische Umsetzung und Integration in die IT-Umgebung des UBA

Da ohne einen abgestimmten Ansatz zur technischen Umsetzung und Integration die Implementation jedes Fachmodells in eine Anwendung obsolet ist, beginnt das Konzept mit diesem Teil.

Für die Analyse und Festlegung der technischen und organisatorischen Vorgehensweise sind verschiedene Quellen genutzt worden:

- ▶ Leistungsbeschreibung des UBA
- ▶ Abstimmung mit der Projektleitung UBA
- ▶ Abstimmung zwischen Projektleitung und Rechenzentrum UBA
- ▶ Architekturrichtlinie IT-Bund (ARIT)

Diese legen folgende Dinge fest:

- ▶ Wo das System betrieben wird

- ▶ Ob und wie das System unterteilt wird
- ▶ In welchen Netzen dann welche Teile betrieben werden
- ▶ Auf welchen Betriebssystemen die Teile des Systems betrieben werden
- ▶ Wie die Ausstattung des Anwendungsservers und des Datenbankservers initial ist
- ▶ Wie sich die Forderung des IT-NetzG⁵ auswirkt
- ▶ Auf Basis welcher Protokolle die Teile des Systems miteinander kommunizieren und welche weiteren Anforderungen dazu existieren
- ▶ Welche Datenbank eingesetzt wird
- ▶ Welche Frameworks für die Entwicklung eingesetzt werden
- ▶ Wie der Zugriff der Auftragnehmerin während der Testphase und im Produktivbetrieb erfolgt
- ▶ Wann die Anwendung verfügbar sein soll

Folgende Anforderungen an das System sind festgelegt, erarbeitet bzw. abgestimmt worden:

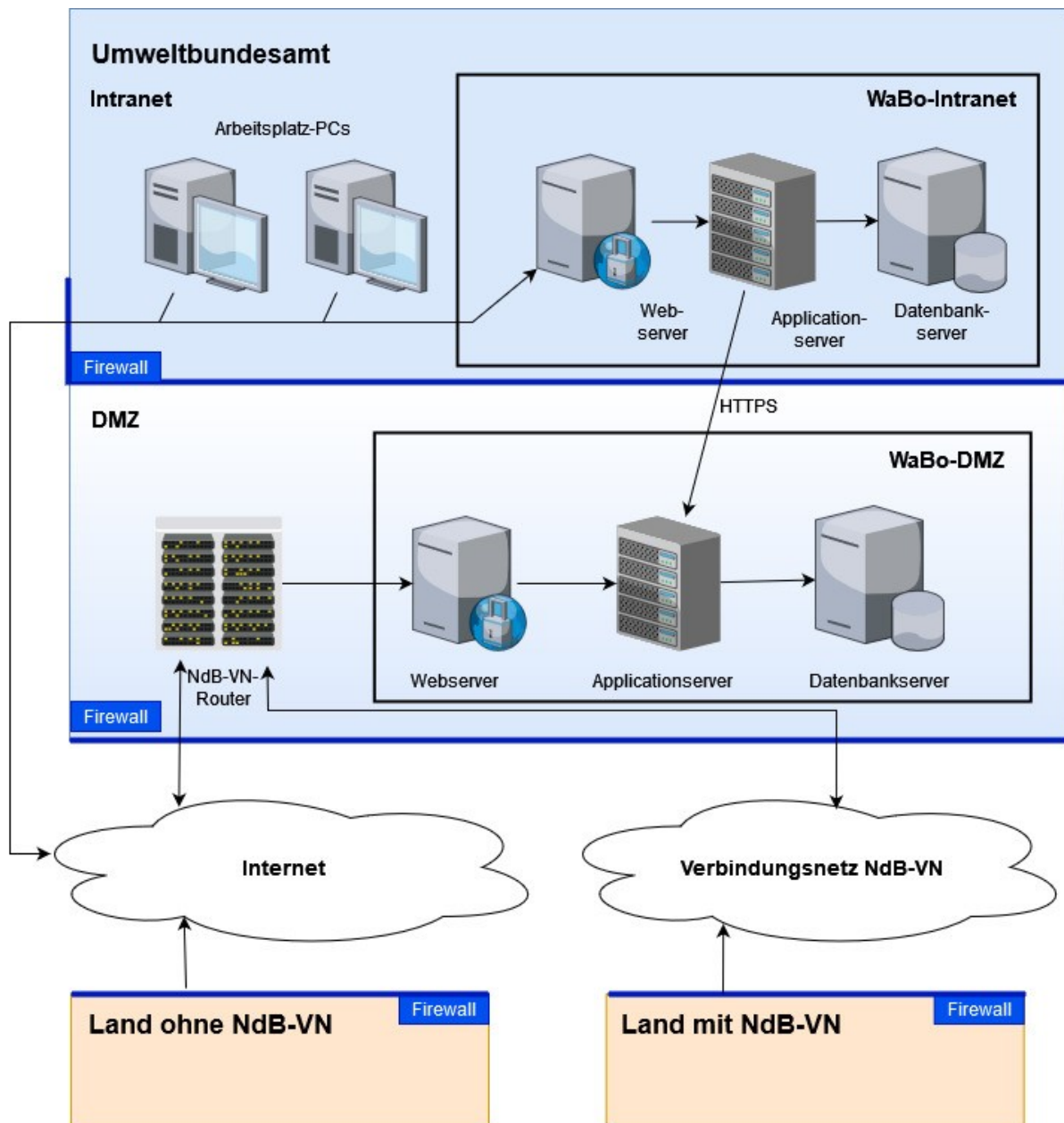
- ▶ Das System wird im UBA betrieben und es wird ein Testsystem im UBA installiert und betrieben, das erst wenn es fertig gestellt ist, auf das Produktivsystem umgestellt wird
- ▶ Weiterhin wird das System unterteilt in
 - Ein System in der UBA DMZ für den Zugang zu den IT-NetzG-konformen Netzen und die Außenkommunikation mit den Datenlieferanten (DMZ-System) und
 - Ein System im UBA Intranet für die Innenkommunikation zu den Nutzern im UBA (Intranet-System)
- ▶ Diese Systeme werden nicht weiter unterteilt, sondern es werden die Teile Webserver, Application Server und Datenbankserver auf einem einzigen System installiert
 - Weiterhin ist eine Unterteilung in IT-NetzG-Dienste und nicht-IT-NetzG-Dienste (in der DMZ) nicht notwendig, da die IT-NetzG-Konformität durch das korrekte Routen von Netzwerkpaketen vom und zum DMZ-System erreicht werden kann und sollte
- ▶ Die Kommunikation zwischen dem DMZ-System und dem Intranet-System erfolgt ausschließlich über HTTPS
 - Eine Kommunikation kann nur vom Intranet aus initiiert werden
- ▶ Die Systeme werden unter Ubuntu 22.04 betrieben
- ▶ Die Server werden jeweils initial als virtuelle Maschine mit 15 GB Festplattenspeicher, 4 GB RAM, nginx Webserver, uWSGI Application Server und PostgreSQL Datenbankserver ausgestattet

⁵ Gesetzliche Grundlage für das Verbindungsnetz NdB-VN zwischen Bund und Ländern

- UBA kann unter Windows als Webserver Apache und Microsoft IIS einsetzen und administrieren. Unter Linux bestehen keine Beschränkungen.
 - Die Auftragnehmerin kann als Webserver Apache und nginx einsetzen und administrieren. Beide sind in der aktuellen Fassung der Architekturrichtlinie IT-Bund als SOLL-Kriterium (TIAS-06/ID: AS-9005-R02) aufgeführt. Es wurde zugunsten des leichtgewichtigeren nginx entschieden.
 - Als Frameworks für die Entwicklung werden Python/Django im Backend und Vue.js für das Frontend eingesetzt
 - Als relationales Datenbankmanagementsystem (RDBMS) kommt PostgreSQL in aktueller Version (≥ 13.4) zum Einsatz
 - Als GIS-Datenbank-Framework wird PostGIS in aktueller Version (≥ 3.1) eingesetzt
- Der Zugriff auf Test- und Produktivsysteme erfolgt zukünftig über die Genugate Appliance
- Die Verfügbarkeit der Anwendung ist im normalen Tagesbetrieb von 09:00 - 17:00 Uhr wichtig

Aus den vorgenannten Anforderungen wurde eine IT-Architektur entwickelt, deren grafische Darstellung sich in Abbildung 1 findet.

Abbildung 1: FIS Wasser-Boden IT-Architektur



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Die Anforderungen bzgl. der Aufteilung der Anwendungen, dem Kommunikationsprotokoll über die Grenze zwischen UBA Intranet und DMZ hinweg sowie der einzusetzenden Betriebssysteme und der Ausstattung der Systeme sind durch die organisatorischen Belange des UBA RZ ebenso vorgegeben wie die Entscheidung, Webserver, Anwendungsserver (Application Server) und Datenbankserver jeweils gemeinsam auf einem System zu betreiben.

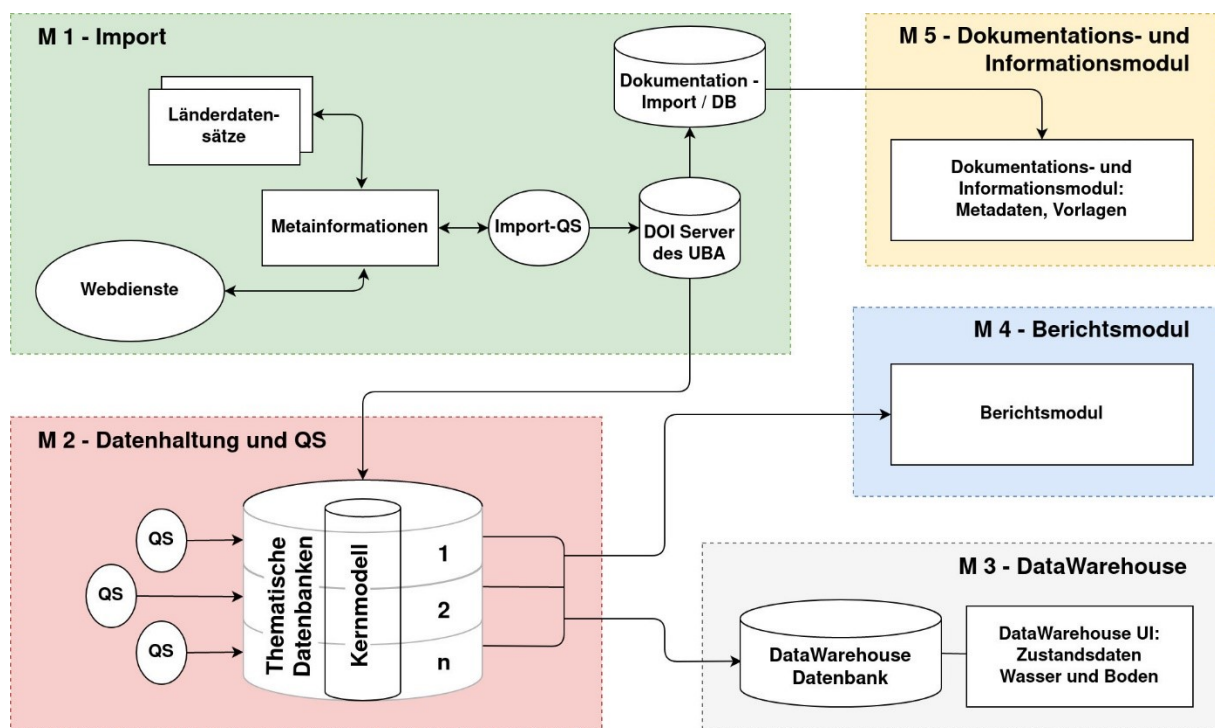
Andere Anforderungen sind durch die Architekturrichtlinie IT Bund festgelegt. Die Festlegung auf Python/Django und nicht auf Java und eines der zahlreichen Frameworks entspringen der Anforderung, das FIS langfristig lizenzkostenfrei betreiben zu können und kostengünstig pflegbar zu halten. Eine durch Millionen von Nutzern verursachte Rechenlast, die den Einsatz von Java notwendig machen würde, kann sicher ausgeschlossen werden. Gleichzeitig hat sich Python als Sprachplattform für die Auswertung großer Datenmengen etabliert. Da der Markt der unterschiedlichen Frameworks für Java sehr breit gefächert ist und sich IT-Unternehmen generell auf

bestimmte Frameworks spezialisieren, ist die Menge der Anbieter mit Erfahrung in Python/Django inzwischen größer. Die Erfahrung bei der Auftragnehmerin sowohl mit Java als auch mit Python zeigt, dass Ergebnisse mit gleichem Qualitätsniveau unter Einsatz von Python mit geringerem Aufwand erreicht werden.

4.2 Module der Anwendung

Der Entwurf des konzeptionellen Modells für das FIS Wasser-Boden wurde insbesondere im Hinblick auf die Datenhaltung (Modul M2) überarbeitet.

Abbildung 2: Module der Anwendung



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Um es zu ermöglichen, Messwerte Umweltkompartiment-übergreifend abzufragen, ist eine gemeinsame Datenhaltung in einem für alle Umweltkompartimente genutzten Kernmodell notwendig. Dies ist auch aus IT-technischen Gründen sinnvoll, da die Bildung von Doppelstrukturen möglichst verhindert werden soll (Leser & Naumann, 2006). Weitere Erläuterungen dazu finden sich im Kap. 11.2, Datenmodell WaBo-Intranet.

Die Verwaltung der Inhalte des DataWarehouse (M 3) ist in die Datenbank für die Umweltdaten-Datenhaltung (M 2) integriert worden. Die Ablage der Inhalte des DataWarehouse erfolgt im Dateisystem.

5 Aufnahme der Nutzungsprozesse

Um eine für jedes Umweltkompartiment gleiche Vorgehensweise bei der Prozessaufnahme zu gewährleisten, wurde zu Beginn des Projekts ein Fragenkatalog (siehe Anhang A) entwickelt, der den gesamten Umfang der für die Dokumentation notwendigen Themen abdeckt und der für alle vier Umweltmedien angewandt werden kann.

5.1 Fließgewässer

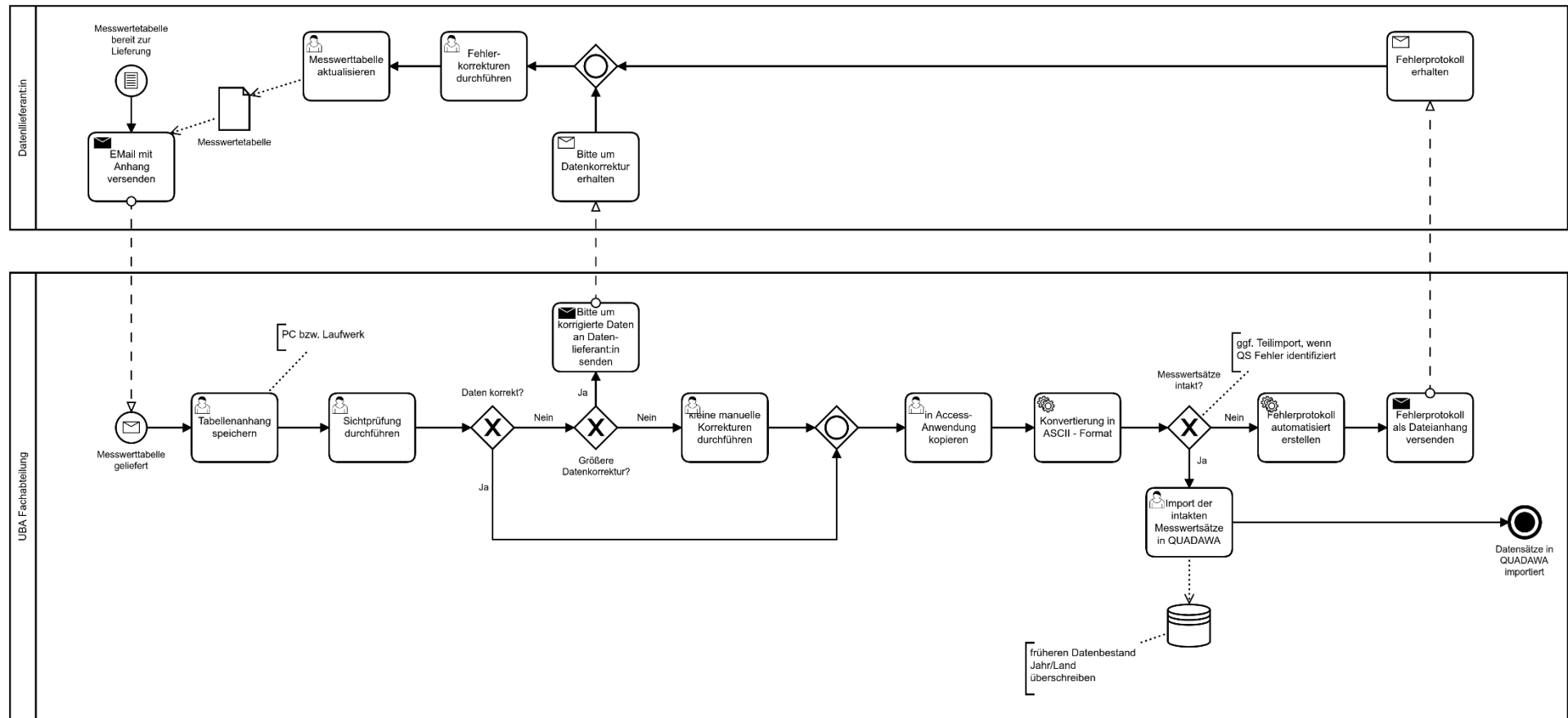
Die Prozess- und Modellaufnahme für das Umweltkompartiment Fließgewässer wurde in Zusammenarbeit mit dem UBA Fachgebiet II 2.4, Binnengewässer durchgeführt. Das Dokument zur Prozessanalyse Fließgewässer befindet sich im Anhang B.

Die aus der Prozesserhebung und -analyse erlangten Erkenntnisse reichten in Verbindung mit den bereitgestellten Unterlagen aus, um einen Prozess im FIS Wasser-Boden zu realisieren, der den vorhandenen Funktionsumfang des bestehenden Prozesses ersetzen konnte.

Besonderes Augenmerk wurde daraufgelegt, dass das FIS für Fließgewässer die in der Prozessaufnahme (Anhang B.9, „Prozessanalyse Fließgewässer, Verschiedenes“) aufgeführte Möglichkeit, Abfragen auf z. B. MS-Access auszuführen (und daraus Excel-Tabellen zu exportieren), adäquat ersetzt, um die bisherigen Verarbeitungsmöglichkeiten weiter zu bieten.

Der Ist-Prozess der Datenlieferung Fließgewässer per .xlsx-Datei stellte sich wie folgt dar:

Abbildung 3: Ist-Prozess Datenlieferung Fließgewässer



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Weiterhin war beim Umweltkompartiment Fließgewässer eine Datenlieferung per ASCII-Festbreitenformat bei etwa der Hälfte der Datenlieferanten etabliert. Da dieses sinnvoll nur maschinell erstellt werden kann, ist es gegenüber Excel-Dateien weniger fehleranfällig. Weiterhin entfällt bei der bislang genutzten Softwarelandschaft die Notwendigkeit der Konvertierung (in obigem Diagramm zwei Prozessschritte) in das ASCII-Festbreitenformat.

Festbreitenformate gelten jedoch generell als veraltet und finden in der Architekturrichtlinie⁶ keine Erwähnung. Als gemeinsamer Nenner der Datenliefernden aller Kompartimente wurden tabellenbasierte Formate (.xlsx und CSV) identifiziert.

5.2 Seen und Grundwasser

Die Ist-Prozesse zur Aufnahme von Messwerten für die Umweltkompartimente Seen und Grundwasser unterschieden sich vom Ist-Prozess der Fließgewässer lediglich in zwei Punkten:

- ▶ Es entfiel das Kopieren in eine (kleine) Access-Datenbankanwendung und die Konvertierung in ein direkt importierbares Format
- ▶ Es wurden keine Teildatenimporte durchgeführt, sondern die Gesamtlieferung oder kein Datensatz.

5.3 Boden

Für Bodendaten existierten zwei Importprozesse. Der erste Prozess behandelte Datenlieferungen der bayerischen Ämter LfL (Landesanstalt für Landwirtschaft), LfU (Bayerisches Geologisches Landesamt, München und Bayerisches Landesamt für Umwelt - Geologischer Dienst), und LWF (Landesanstalt für Wald und Forstwirtschaft).

Der zweite Prozess auf Basis wechselnder Formate und Strukturen wurde durch einen einheitlichen Importansatz auf Basis tabellenbasierter Formate (.xlsx, CSV) analog zu den Wasserkompartimenten, ersetzt. Das entsprechende Nachrichtenmodell ist in Kap. 12.6, „Boden“ beschrieben.

⁶ Die Architekturrichtlinie IT-Bund sieht in IDAS-03, „Darstellung von strukturierten Daten zum Zweck des Datenaustauschs“ die Formate XML, JSON oder YAML vor. Sollen Dateien ausgetauscht (und verarbeitet) werden, so ist noch IDAS-11 „Nutzung von einheitlichen Austauschformaten für Dokumente, die in Bearbeitung sind“ mit den Formaten Office Open XML (OOXML) oder Open Document Format for Office Applications 1.2 (ODF) sowie Comma-Separated Values (CSV) relevant.

6 Zielprozesse und Übergangsprozess Datenlieferung

Die hier beschriebenen Zielprozesse stellen eine sehr weitgehende Ausbaustufe dar, die in agilen Fachabstimmungen iterativ an die Bedürfnisse der beteiligten Fachabteilungen angepasst wurden. Dabei wurde von der Vorstellung, in größerem Umfang Datenlieferungen im XML-Format zu erhalten, Abstand genommen. Hintergrund ist die Tatsache, dass genau ein Datenlieferant Daten über eine REST-Schnittstelle bereitstellt und dabei eine XML-Struktur vorgibt.

Der Zielprozess ist technikunabhängig und verändert sich daher nicht, egal ob die Nachrichtenübermittlung in der Übergangsphase manuell mit dem Webbrowser erfolgt oder ob ein automatisierter Austausch von elektronischen Nachrichten zwischen Systemen⁷ erfolgt.

6.1 Zielprozess Datenlieferung

Der neue Prozess der Datenlieferung hat die Verfahrensbeteiligten:

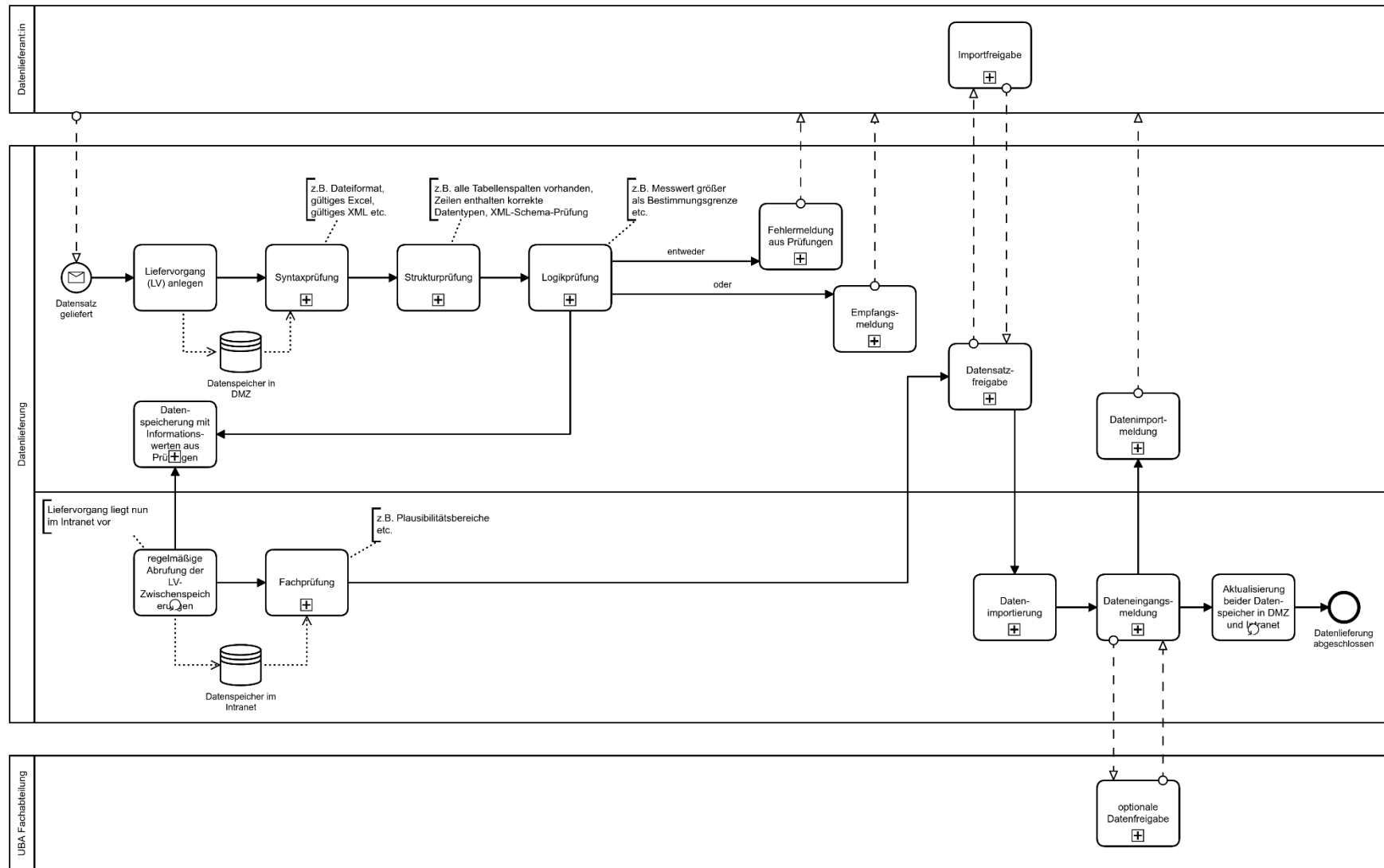
- ▶ Datenlieferant:in
- ▶ UBA DMZ
- ▶ UBA Intranet
- ▶ UBA Fachabteilung (unabhängig vom Umweltkompartiment) und
- ▶ Verfahrensherstellerin (ENDA)

Er sieht gestufte Datenprüfungen und Rückmeldeschleifen vor und ist weitgehend automatisiert — benötigt folglich nur an den Stellen manuelle Eingriffe, an denen dies nicht automatisiert werden kann, z. B. im Fall einer Freigabe.

Der Zielprozess für die Datenlieferung — zunächst aus strategischer Sicht — stellt sich wie folgt dar:

⁷ Der automatisierte Austausch von Nachrichten zwischen elektronischen Systemen wird heutzutage (gem. Architekturkonzept IT-Bund) mit Hilfe von RESTful Webservices und festgelegten Nachrichtenformaten in XML (oder JSON) realisiert. Dabei erfolgt der Datentransport unter Einsatz von zertifikatbasiertem HTTPS zur Sicherung gegen Manipulationen und auch gegen das Mitlesen durch nicht autorisierte Dritte. Vor der Ausführung von Nutz-Operationen muss eine erfolgreiche Authentifizierung stattfinden, um die Nutzung der Schnittstelle durch nicht autorisierte Dritte zu verhindern. Da bislang keine technischen Beschreibungen über Schnittstellen von Datenlieferanten vorliegen, werden hier keine fachlich weitergehenden Festlegungen gemacht.

Abbildung 4: Zielprozess Datenlieferung (strategische Sicht)



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Die Zahl der versandten Benachrichtigungen des Prozesses stellt den Umfang dar, der noch als sinnvoll betrachtet werden kann. Die notwendig und geforderte Dokumentation aller Datenimporte ist Teil des Prozesses.

Der in BPMN2 modellierte Prozess ist weitgehend selbsterklärend — dafür sind Prozessmodellierungsnotationen als Bildsprache zur Darstellung und zum leichten Nachvollziehen der Schrittabfolge, aus denen Prozesse bestehen, entwickelt worden. Einige Hinweise sind jedoch sinnvoll.

Mit dem Eintreffen eines vom Datenlieferanten hochgeladenen Datensatzes wird ein Liefervorgang (LV) angelegt und in verschiedenen weiteren Prozessschritten um Informationen ergänzt.

Der Lieferprozess sieht vor, dass nach dem erfolgreichen Durchlaufen der inhaltlichen und Plausibilitäts-Prüfungen eine Datensatzfreigabe bei der Lieferant:in angefordert wird und erst nach deren Freigabe der Import durchgeführt wird. Auftraggeber und Auftragnehmerin sind übereingekommen, auf eine Freigabe durch die Datenlieferant:innen zu verzichten. Dies ist sinnvoll, da die Importstatistik⁸ erst nach dem Import erstellt wird und dem/der Lieferant:in daher vor allem Meldungen zur Fehlerfreiheit vorliegen.

Auch auf eine Freigabe durch das UBA wird verzichtet, da die Datenlieferungen nur durch autorisierte Fachbetreuende erfolgen können und das UBA von qualitätsgesicherten Eingangsdaten aus den Datenhaltungen der Bundesländer ausgehen kann.

Mit dem erfolgreichen Import in das FIS und der Dokumentation des Liefervorgangs in der Datenbank des FIS wird daraus ein die Lieferung vollständig beschreibender Metadatensatz.

Ein alternatives Prozessszenario könnte vorsehen, für jedes Umweltkompartiment individuell entweder manuelle oder automatische Freigabe festlegen zu können. Die Änderung des Verhaltens würde dann über eine Konfiguration erfolgen.

Von den derzeit vier Erfolgsmeldungstypen

- ▶ Empfangsmeldung
- ▶ Datensatzfreigabeanforderung Lieferant:in
- ▶ Dateneingangsmeldung und
- ▶ Freigabemeldung UBA

sind Empfangsmeldung und Dateneingangsmeldung unverzichtbar.

Der im FIS realisierte Zielprozess ohne Freigaben kann auch für die Lieferung von Bodendaten von Untersuchungsflächen eingesetzt werden. Entsprechende Nachrichtenmodelle sind für alle Umweltkompartimente entwickelt worden.

Für die Lieferung von Messwerten für das Umweltkompartiment Fließgewässer besteht die Anforderung, ausschließlich solche Prüfungen auszuführen, die notwendig sind, um die Daten in der Datenbank wiederauffindbar abzulegen (technische Verarbeitbarkeit). Dies entspricht vollständig dem bisherigen Prüfungsumfang der Fließgewässer-Datenbankanwendung QUADAWA. Darüber hinaus gehende Prüfungen wie z. B. Konzentrationsbereiche sollen nicht umgesetzt werden. Daher bleiben die entsprechenden Prozessschritte „Daten (LV) logisch prüfen“ (im

⁸ Die Importstatistik soll die Gesamtzahl der importierten Messwerte, und der Zahl der Messwerte je Messgröße (Parameter), je Bundesland und je Jahr enthalten. Kombinationen dieser Kriterien sind nicht vorgesehen.

Bereich UBA DMZ) und „Daten LV fachlich prüfen“ (im Bereich UBA Intranet) in diesem Fall ohne Inhalt.

In jedem Fall werden Datenlieferungen nur vollständig importiert oder gar nicht⁹. Sind einzelne Messwerte nicht verarbeitbar oder nach den jeweiligen QS-Prüfungen fehlerhaft, so wird der Datensatz mit Fehlerprotokoll automatisiert zurückgewiesen, ohne dass UBA davon in Kenntnis gesetzt wird. Die Datenlieferant:in hat die Möglichkeit, die Importdatei zu korrigieren und einen erneuten Importversuch zu starten.

6.2 Zielprozess Produkterstellung

Für die Erfüllung verschiedener Berichtspflichten sind aus den im FIS Wasser-Boden vorhandenen Daten Berichte, z. B. an die EUA, zu generieren. Die unterstützten Formate (Excel97 .xls und XML) und die dabei einzuhaltenden Strukturen sind jeweils festgelegt und unterliegen i.d.R. nur geringen Veränderungen von Berichtsperiode zu Berichtsperiode. Sie resultieren meist aus Änderungen am strukturellen Umfang der zu liefernden Berichtsdaten.

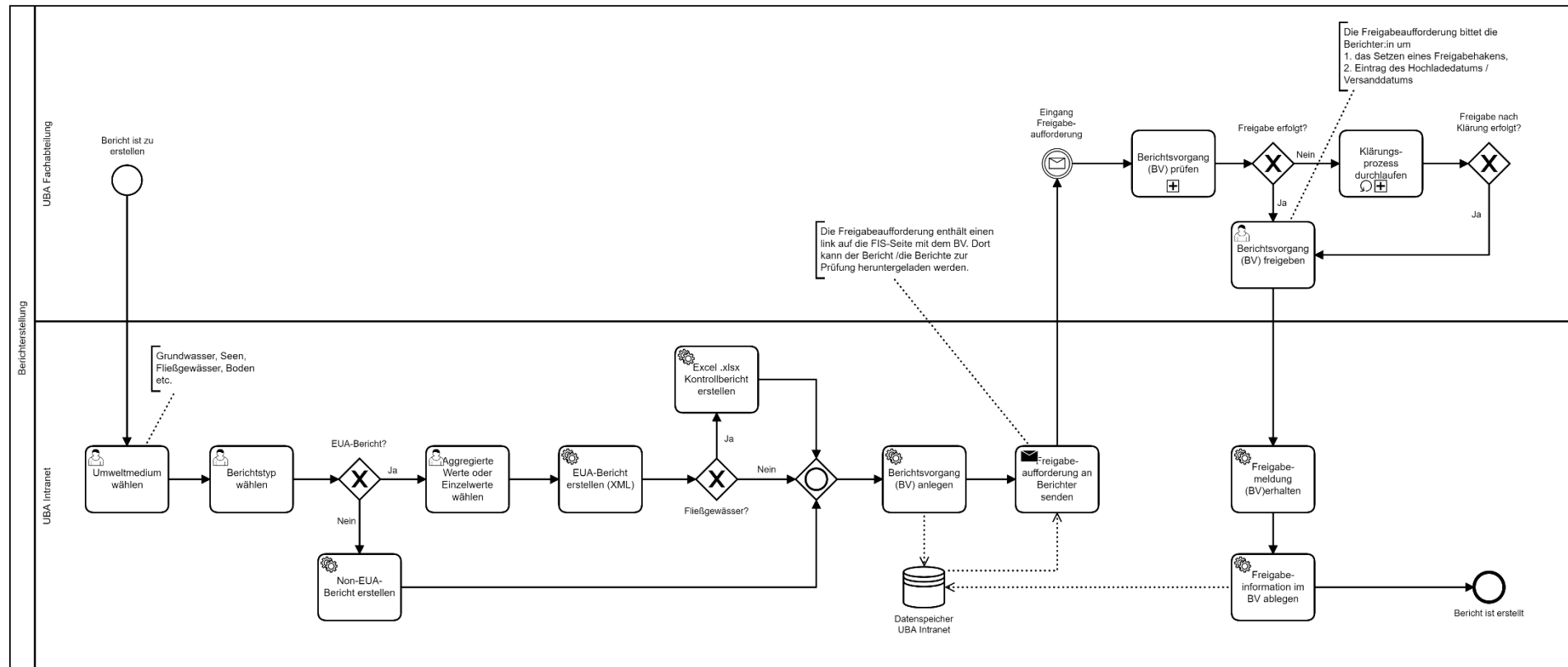
Da in den Umweltmedien, für die solche Berichtsverpflichtungen vorliegen, die Kriterien zur Auswahl der Daten gleich sind, ergeben sich auch keine Abweichungen für den Prozess. Folgende Kriterien zur Auswahl der Daten sind notwendig und sinnvoll:

- ▶ Berichtszeitraum mit Anfangs- und Enddatum
- ▶ Parameterumfang (oder Messgrößenumfang)
- ▶ Messnetz
- ▶ Bundesländer (optional, Boden)

Der Zielprozess zur Produkterstellung (Berichterstellung) ist nachfolgend abgebildet und zur besseren Lesbarkeit auch im separaten ZIP-Container **Berichtsgrafiken.zip** zum Abschlussbericht hinterlegt:

⁹ Auch wenn von 100.000 Zeilen nur eine nicht verarbeitbar oder inhaltlich fehlerhaft ist, wird der Datensatz zurückgewiesen. Inhaltliche Prüfungen finden beim Import von Fließgewässerdaten nicht statt. Selbstverständlich werden bei mehreren Fehlern alle berichtet, damit nur ein Korrekturvorgang notwendig ist.

Abbildung 6: Zielprozess Produkterstellung



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

6.3 Übergangsprozess Datenlieferung

Dieser Prozess liefert die gleichen Ergebnisse, wie der Zielprozess zur Datenlieferung.

Der Unterschied liegt darin, dass tabellenorientierte (auf .xlsx-Dateien basierende) Datenlieferungen i.d.R. manuell zusammengestellt werden, wohingegen XML-Dateien von einem Programm automatisiert erstellt werden. Ist ein Programm zur Erstellung von XML-Dateien fertig entwickelt und getestet, ist diese Art der Datenbereitstellung weniger fehleranfällig.

Die Erfahrung mit dem Austausch tabellenbasierter Daten hingegen zeigen, dass folgende (die Liste ist nicht abschließend) Fehler immer wieder passieren:

- ▶ Die Tabellenstruktur wird verändert
- ▶ Die Spaltenköpfe werden anders benannt, als vereinbart
- ▶ Listenwerte entsprechen nicht einem erwarteten Wert (z. B. Parameter unbekannt, Einheit unbekannt etc.)
- ▶ Bedingungen i.S.v. „wenn der Messwerte unterhalb der BG liegt, muss die BG angegeben und das entsprechende Kennzeichen gesetzt werden“ und andere bedingte Erstellungsvorschriften werden nicht eingehalten
- ▶ Es werden Werte aus Tabellen mit Werten in mg/kg mit Werten in µg/kg in einer Spalte zusammenkopiert
- ▶ Es werden für Nitrat oder Ortho-Phosphat die Werte für Nitrat-Stickstoff oder Ortho-Phosphat-Phosphor angegeben
- ▶ Es wird bei Leitfähigkeitswerten ein Wert in mS/m ausgewiesen, aber in µS/cm geliefert

Folglich sind die bislang manuell durchgeführten Sichtprüfungen soweit möglich und wirtschaftlich vertretbar/sinnvoll automatisiert worden und es wird eine intensivere Nachkontrolle bei UBA erfolgen, bevor eine Datenlieferung freigegeben werden kann.

Ziel auch dieses Übergangsprozesses ist es, dass Datenlieferungen mit einem geringeren Aufwand als bislang entgegengenommen werden können und dass eine exakte Dokumentation von Datenlieferant, Lieferzeitpunkt, Lieferungsinhalt und Prüfungsergebnissen erstellt, abrufbar hinterlegt und mit jedem ggf. importierten Wert verknüpft wird.

Ein weiterer Unterschied liegt darin, dass die Prüfungen der Syntax, also der Verarbeitbarkeit der gelieferten Daten, die bei XML-Dateien ohne nennenswerten Aufwand durchgeführt werden können, je Nachrichtenmodell und somit je Umweltkompartiment (und bei Boden für zwei Nachrichtenmodelle) recht aufwändig programmiert werden müssen.

6.4 Alternativer Zielprozess zur Datenakquisition

Das Landesamt für Umwelt, Naturschutz und Geologie Mecklenburg-Vorpommern (LUNG MV) hat eine REST-Schnittstelle für den Abruf von Grundwasserdaten durch das Umweltbundsamt mit dem System FIS Wasser-Boden eingerichtet und zur Produktionsreife weiterentwickelt. In der Anwendung kann über den Menüpunkt „Länder-API Datenabfragen“ diese Schnittstelle angesprochen werden. Dabei werden die Daten automatisiert abgerufen, für den Import aufbereitet und grundlegenden Prüfungen (z. B. Existenz der Messstelle im FIS Wasser-Boden) unterzogen. Nachfolgend wird automatisch der Importprozess gestartet.

Das FIS Wasser-Boden sieht vor, dass sowohl für die anderen Bundesländer also auch für die anderen Umweltkompartimente solche Schnittstellen für die Datenakquisition genutzt werden können. Zukünftig können Implementationsaufwände auf der Seite des FIS Wasser-Boden reduziert werden, wenn die Länder die Struktur ihrer Schnittstellen technisch und funktional einheitlich gestalten.

7 Konzept Anwendung

Die Anwendung wurde als Webanwendung entwickelt und ist in den gängigen Webbrowsern (Google Chrome, Mozilla Firefox, Apple Safari sowie der aktuellen Version von Microsoft Edge) vollumfänglich lauffähig.

Nutzer:innen können sich unter Angabe ihrer Emailadresse und eines Passworts anmelden. Die Konten jeder Nutzer:in sind an eine oder mehrere Nutzer:innen:gruppen gebunden, die einerseits den Rechteumfang beschreiben und andererseits von der Anwendung genutzt werden, um die Zugehörigkeit zu einem oder mehreren Umweltkompartimenten zu prüfen.

Die Zugehörigkeit zu einem oder mehreren Umweltkompartimenten beschränkt den Zugang zu kompartimentspezifischen Anwendungsteilen wie der nur für Fließgewässer relevanten Aggregation und filtert an gemeinsam genutzten Anwendungsteilen die Menge der angezeigten Datensätze. Dies ist insbesondere für die nachfolgend beschriebenen Funktionalitäten von Parametergruppen und Messnetzen sinnvoll, um nur auf die eigenen Fachdaten Zugriff zu haben.

7.1 Anzeige von Daten

Daten werden generell tabellarisch dargestellt, wenn mehrere Datensätze visualisiert werden. Aus der tabellarischen Sicht können einzelne Sätze in einem Formularmodus angezeigt, bearbeitet und gespeichert werden.

Die tabellarische Darstellung von Messstellen zeigt nur die Messstellen des gewählten Umweltkompartiments – bei Nutzenden mit nur einem Umweltkompartiment demzufolge nur die Messstellen dieses einen Umweltkompartiments, das automatisch durch Login und Gruppenzugehörigkeit ermittelt wird. Weiterhin existiert dort die Möglichkeit der Filterung nach:

- ▶ Messnetzen (siehe unten)
- ▶ Zeichenkette (mit inkrementeller Filterung bei der Eingabe), die im Namen, in der Messstellennummer und im Bundeslandnamen gesucht wird

Die tabellarische Darstellung von Parameter-Aliassen zeigt nur jene Parameter, zu denen im jeweiligen Umweltkompartiment ein sog. Parameter-Alias existiert. Beispielsweise bekommen Nutzende der Wasserkompartimente keine bodenspezifischen Feldkapazitätsparameter angezeigt. Daneben existieren Filter nach:

- ▶ Parametergruppen (siehe unten)
- ▶ Zeichenkette (mit inkrementeller Filterung bei der Eingabe), die im Namen, in der physikalischen Größe, der Matrix und bei Wasserkompartimenten in der LAWA-Nummer gesucht wird
- ▶ Physikalischer Größe (Temperatur, Massendichte, Leitfähigkeit etc.)
- ▶ Matrix (Gesamtgehalt, gelöste Konzentration etc.)
- ▶ Messgrößentyp (z. B. Summenparameter, Stoff, physikalische Größe)

Das Formular kann per Klick auf die Darstellung aller Parameter umgestellt werden. Die o.g. Filter können dann weiterhin genutzt werden. Um für ein Umweltkompartiment einen neuen Parameter-Alias anzulegen oder einen neuen Parameter einzuführen, müssen Nutzende über entsprechende Verwaltungsrechte verfügen.

7.2 Funktionalität zur Nutzung von Parametergruppen

Die Anwendung erlaubt es bei der Abfrage von Parametern und zugehörigen Messwerten, Parametergruppen zur Filterung der Ergebnismenge einzusetzen. Eine Parametergruppe besitzt einen Namen, eine Nutzergruppe (i.S.v. Fließgewässer-Nutzer:in, Boden-Nutzer:in, ...) und eine manuell konfigurierbare Menge zugeordneter Parameter aus der kompartmentübergreifenden Hauptliste der Parameter.

Die Nutzergruppe ist notwendig, um den Bearbeiter:innen der jeweiligen Umweltkompartimente eine Sicht auf nur ihre Parametergruppen zu ermöglichen.

Hintergrund der Einführung von Parametergruppen ist, dass historisch gewachsene Konstrukte wie in Grundwasser und Fließgewässern mit unterschiedlichen Parametern besetzte Gruppen und Obergruppen nicht zu einem Flickwerk im Datenmodell der FIS Wasser-Boden-Datenbank und deren Nutzungsoberfläche führen sollen.

Weiterhin ist eine Möglichkeit zur reproduzierbaren Filterung nach aufgabenspezifisch zusammengestellten Parametergruppen ein mächtiges Werkzeug, das Bestandteil eines modernen FIS sein sollte, um die Bearbeitungsprozesse gut zu unterstützen.

Parametergruppen können von bestehenden Parametergruppen abgeleitet werden. Sie erhalten dann als initiale Füllung alle Parameter der Gruppe, aus der sie abgeleitet wurden. Dies kann auch mehrfach durchgeführt werden, um die Parameter weiterer Gruppen hinzuzufügen.

Innerhalb der Verwaltungsoberfläche für Parametergruppen ist als Filtermöglichkeit die inkrementelle Einschränkung nach dem Anfang oder einem Teil des Parameternamens implementiert.

Um die Komplexität der Verwaltung von Parametersätzen v.a. durch die Nutzenden nicht zu hoch werden zu lassen, wird keine mehrstufige Verwaltung vorgesehen. Sollte eine Strukturierung notwendig sein, soll diese durch entsprechende Namensgebung erreicht werden.

Die hier beschriebenen Parametergruppen sind für die tägliche Arbeit mit dem Abfragewerkzeug des FIS Wasser-Boden gedacht. Die für die Erstellung von Berichten notwendigen Parameternamen werden getrennt gehalten und verwaltet.

7.3 Funktionalität zur Nutzung von Messnetzen

Analog zu den Parametergruppen wurde eine Möglichkeit zur Erstellung und Festlegung von Messnetzen (benannte Mengen von Messstellen)¹⁰ geschaffen. Primärer Nutzen von Messnetzen ist deren Verwendung im Abfragewerkzeug. Auf diese Weise kann mit geringem Aufwand eine Abfrage erstellt werden, deren Ergebnisse nur aus einem Messnetz stammen. Weiterhin ist diese Methode sehr wenig fehleranfällig gegenüber dem Erstellen einer Abfrage, in der alle zu berücksichtigenden Messstellen einzeln angegeben werden.

Die Zusammenstellung eines Messnetzes soll wieder analog zu der der Parametergruppen erfolgen können, wobei die Filterung auf LAWA-Nummern entfällt und durch die Möglichkeit der Auswahl eines oder mehrerer Bundesländer ersetzt wird.

Weiterhin ist eine Möglichkeit zur reproduzierbaren Filterung nach aufgabenspezifisch zusammengestellten Messnetzen ein mächtiges Werkzeug, das Bestandteil eines modernen FIS sein sollte, um die Bearbeitungsprozesse gut zu unterstützen.

¹⁰ Dies meint auch die Untersuchungsflächen des Umweltkompartiments Boden

Zur Unterstützung der Bearbeitungsprozesse sind die Filtermöglichkeiten nach Parametergruppen und Messnetzen kombinierbar. Hier wird besonderes Augenmerk auf die Entwicklung des Anwendungsdesigns gelegt, um die Nutzung einfach und möglichst selbsterklärend zu halten.

Die Messnetze werden direkt vom Berichterstellungsteil der Anwendung für die automatisierte Erstellung entsprechender Berichte genutzt. Sie werden daher jeweils mit einer Berichtspflicht verknüpfbar sein, jedoch nicht verknüpft sein müssen.

7.4 Fließgewässer Datenzugriff

Die Bearbeitung der Fließgewässerdaten erforderte ursprünglich manuelle Kontrollen und Aufbereitungsschritte, die mit Access und Excel durchgeführt wurden. Um diese Möglichkeit auch mit FIS Wasser-Boden zu gewährleisten, wurde ein Zugriff auf die Fließgewässerdaten aus Access per ODBC-Schnittstelle eingerichtet, so dass FG II 2.4 lesenden Zugriff auf Messwerte und Messstellen und die fließgewässerspezifischen Parameter sowie alle weiteren, auch in der Ursprungsdatenbank QUADAWA vorhandenen Tabellen erhält.

7.5 Seen – Modell und Anwendungsnutzung

Seen werden hier explizit angesprochen, da das Datenmodell der Seenmessstellen verändert wurde. Dabei geht es um eine Vielzahl von Attributen zu verschiedenen Landnutzungsarten, die jeweils in km² und % im Einzugsgebiet mit Werten belegt werden konnten. Die Werte dieser Attribute sind ständig in Gefahr, veraltet zu sein. Weiterhin ist die Abbildung von solchen Attribut-sammlungen im Datenmodell der Anwendung unschön, glücklicherweise jedoch auch vermeidbar.

Die Angaben der verschiedenen Kategorien von Landnutzungsflächen in km² und % im Einzugsgebiet werden durch Verschneiden des See-einzugsgebiets¹¹ mit den Landnutzungsflächen der aktuellen Ausgabe von Corine Land Cover erzeugt. Dazu wurde die Kategoriespezifikation von Corine Land Cover¹² ausgewertet und für die derzeit verwendeten Kategorien ein Mapping auf die Kategorien von Corine durchgeführt.

Die folgenden Zuordnungen konnten alle derzeitigen Attribute ablösen:

- ▶ Wald/Forst: 3.1 Forest
- ▶ Gesamte landwirtschaftlich genutzte Fläche: 2. Agricultural areas
- ▶ Grünlandfläche: 2.3 Pastures
- ▶ Ackerfläche: 2.1 Arable land (ggf. mit 2.4.1 Annual crops associated with permanent crops; Erst einmal zu andere Flächen)
- ▶ Grünland- + Ackerfläche: Summe der o.g.
- ▶ Urbanisierte Fläche: 1. Artificial surfaces
- ▶ Wasserfläche: 5. Water bodies
- ▶ Andere 3.2 Shrub and/or herbaceous vegetation associations + 3.3 Open spaces with little or no vegetation + 4. Wetlands

¹¹ aus den Geometrien der Wasserkörper (bzw. für Grundwasser: der Grundwasserkörper)

¹² <https://land.copernicus.eu/user-corner/technical-library/corine-land-cover-nomenclature-guidelines/html/>

7.6 Grundwasser

Die Landnutzungsattribute waren in der vorherigen Version der Grundwasseranwendung analog zu den Seen vorhanden. Es wurde vorgesehen, diese analog zu den Seen über die Nutzung der Grundwasserkörpergeometrie und Verschneidung mit Corine Land Cover dynamisch abzuleiten. Es sind dafür die Karten der HÜK250-Charakteristika, der hydrogeologischen Teilräume und der Boden-Klima-Räume in das FIS integriert worden. Daneben wurden die Karten mit den politischen Gebieten Land, Kreis, Regierungsbezirk, Gemeinde integriert.

7.7 Boden

Die Bodendaten sind im Gegensatz zu den Wasserkompartimenten durch sehr umfangreiche Messstellendaten geprägt. Dies betrifft sowohl deren Anzahl (> 46.000) als auch deren strukturellen Umfang, also die Anzahl der Attribute. Die fachliche Unterteilung der Messstellendaten erfolgt bei Boden durch die vier Abteilungen Stammdaten, Aufnahmesituation, Profilkennzeichnung und Horizonte. Dabei ist die Profilkennzeichnung wiederum unterteilt in vier Unterabteilungen und die Horizonte in fünf Unterabteilungen.

Zur Filterung der Messstellen ist eine inkrementelle Suche auf den Namen, den Nummern und den Bundesländern implementiert. Innerhalb der Messstelle kann auf die Messungen per Probenahmedatum und nach Messgröße zugegriffen werden.

7.8 Prüfungen

Im Rahmen der Arbeiten zu den Projekten XUBetrieb (Betriebliche Stamm- und Berichtsdaten) und P23R (Prozessdatenbeschleuniger) wurde eine Kategorisierung von QS-Prüfungen auf Berichtsdaten im Umweltbereich entwickelt.

Da in den o.g. Projekten ausschließlich XML-Dateien und Validierungen per XML-Schema vorgesehen waren, hier jedoch — wenigstens übergangsweise — auch Tabellendaten prüfbar sein sollen, müssen die von den XML-Schema-Validierungen durchgeführten Prüfungen zusätzlich vorgesehen werden. Es ergeben sich folgende Prüfungskategorien, die in Abhängigkeit von der Art der zu importierenden Daten ggf. nicht alle benötigt werden:

► Verarbeitbarkeit

- Bei Tabellendaten
 - Maschinenverarbeitbarkeit prüfen
 - Nur bei CSV: Encoding
 - Nur bei CSV: Parsebarkeit
 - Spaltenköpfe identifizierbar
 - Spaltenköpfe bekannt
 - Wertezellen identifizierbar
 - Bei XML-Dateien
 - Versionsprüfung
 - XML Schema-Validierungen

- ▶ Strukturelle Prüfungen (Erfüllung struktureller Anforderungen)
 - Teil 1: abhängige Pflichtfelder, Listenwerte aus Listen, deren Umfang den Einsatz von XML Schema zur Validierung ausschließt, Wertebereiche
 - Teil 2: Multiplizitäten unter Berücksichtigung der Objektstatus
- ▶ Inhaltliche Prüfungen: Umfang der berichteten Objekte; benötigt Zugriff auf Daten der Vorperiode
- ▶ Plausibilität (logische und fachliche Prüfungen)
 - Objektintern
 - Objektübergreifend
 - Im zeitlichen Verlauf: Zeitreihenuntersuchung; benötigt Zugriff auf Daten möglichst vieler Vorperioden

Ab den inhaltlichen Prüfungen ist der Zugriff auf Objekte der Vorperiode(n) notwendig, weshalb diese vollumfänglich erst im Anwendungsteil des UBA-Intranets durchgeführt werden können. Die mit dem FIS Wasser-Boden abgedeckten Importprozesse sehen keine Prüfung des Umfangs der berichteten Objekte vor, da Lieferungen zum großen Teil inkrementell erfolgen. Als Plausibilitätsprüfungen können Gültigkeitsbereiche für alle Parameter als Warn- oder Fehlergrenzen definiert werden. Diese werden beim Import überprüft.

Da zu allen Umweltkompartimenten nur Daten importiert werden, die die Qualitätsanforderungen erfüllen, können sie im Abfragewerkzeug uneingeschränkt abgefragt werden.

7.9 Editieren von Daten

Es ist implementiert, dass nicht nur Messstellen und weitere zugehörige Stammdaten (z. B. Seen oder Bodenhorizonte) editiert werden können, sondern auch Messwerte. Die vorgenannte Änderungsmöglichkeit bezieht sich nicht auf die zum Messwert gehörende Messstelle und das Datum der Probenahme, da sonst die Messung unvollständig wird und eine neue Messung angelegt werden müsste. Eine Protokollierung im Sinne von „wer hat wann was geändert“, erfolgt nicht.

Ursprünglich sollte die Datenlieferung einen Indikator bekommen, der Änderungen an Messwerten der Lieferung anzeigt. Dahinter stand der Gedanke, für jeden Messwert eine verantwortliche Person benennen zu können oder in den seltenen Fällen manueller Änderungen sehen zu können, dass eine Datenlieferung sich nicht mehr im Originalzustand befindet. In Abstimmungsgesprächen zwischen Auftragnehmerin und UBA trat jedoch zutage, dass für spezifische Aufgabenstellungen gelegentlich adhoc-Datenabfragen bei den Ländern erfolgen, die sich mit dem Inhalt regelmäßiger Datenlieferungen zur Erfüllung von Berichtspflichten überlappen. In diesem Fall wäre die zuerst erfolgte Datenlieferung ebenfalls als geändert zu kennzeichnen. Dies entstellt jedoch den ursprünglichen Sinn des Änderungsindikators, weshalb von der Idee des Änderungsindikators Abstand genommen wurde.

7.10 Abfragewerkzeug

Für den Zugriff auf die Datensätze und die Möglichkeit des Exports wird eingeloggten Benutzern von FIS Wasser-Boden ein Abfragewerkzeug bereitgestellt, das tabellarische Ergebnisse liefert und flexible Filtermöglichkeiten vorsieht.

Das Abfragewerkzeug orientiert sich in Möglichkeiten und Aufbau am für die Bodenanwendung eBIS betriebenen Abfragewerkzeug. Es sind alle vorhandenen Attribute abfragbar, auch im Kompartiment Boden, das für seine sehr zahlreichen Attribute zur Beschreibung des Bodens bekannt ist. Weiterhin können alle Attribute mit Filterbedingungen belegt werden.

Filterkriterien, die Listenattribute betreffen, können durch die Auswahl eines oder mehrerer Listenelemente, die per inkrementeller Suche auch bei langen Listen schnell ausgewählt werden können, gefiltert werden.

Weiterhin sind numerische Felder mit den üblichen numerischen Vergleichsoperatoren und Textfelder mit den üblichen textuellen Vergleichsoperatoren filterbar. Analog gilt dies für Datumsfelder.

Weiterhin ist es möglich, über die selbstdefinierten Parametergruppen und Messnetze zu filtern.

Die Filterkriterien können gruppiert werden, wobei die Gruppen per UND und per ODER verknüpft werden können. Die Filterkriterien innerhalb einer Gruppe können ebenfalls per UND und per ODER verknüpft werden. Es müssen keine Gruppen genutzt werden. Im Normalfall werden einfach Filterbedingungen definiert, die standardmäßig per UND verknüpft werden.

Abfragen können gespeichert und durch Änderung des Namens und erneute Speicherung kopiert werden. Sie sind selbstverständlich änderbar, wobei eine gespeicherte Abfrage allein durch ihre Änderung und ggf. Nutzung nicht versehentlich überschrieben werden kann, sondern explizit gespeichert wird, um die ursprünglich gespeicherte Abfrage zu ändern.

Das Abfragewerkzeug sieht vor, Abfragen in zeitlich frei definierbaren Zyklen regelmäßig automatisiert ausführen zu lassen und das Abfrageergebnis jeweils im Data Warehouse abzulegen. Das Data Warehouse ist ohne Login allen Mitarbeitenden des UBA zugänglich. Somit stellen die zyklischen Abfragen einen Weg dar, Mitarbeitenden anderer Fachgebiete Zugang zu regelmäßig aktualisierten Daten zu ermöglichen.

Die tabellarischen Ergebnismengen sind bis 1.000.000 Datensätze als .xlsx-, .csv- oder .json-Datei herunterladbar, darüber entfällt .xlsx.

Die Fachdaten von FIS Wasser-Boden werden im UBA-Intranet vorgehalten. Da ein Zugriff von außen auf das UBA Intranet aus Sicherheitsgründen ausgeschlossen ist, können Abfrageergebnisse folglich nur aus dem UBA-Intranet abgerufen werden.

7.11 Aggregation von Daten

Die Aggregation von Daten und die Berechnung der Perzentile ist der ursprünglichen Berechnungsroutine der Fließgewässerdatenbank QUADAWA eins zu eins nachempfunden worden. Das Datenmodell von FIS Wasser-Boden ist für die Ablage von aggregierten Daten erweitert worden. Die Ergebnisse können als Tabellendatei im .xlsx-Format exportiert werden.

Die Nutzungsoberfläche der Aggregation kann die Messwertmenge, die zu aggregieren ist, per Jahresbereich (von – bis), Parameter, Matrix, Bundesland, Messstelle, Messart und Probeart festlegen. Mit derselben Funktionalität können die zu exportierenden, aggregierten Daten festgelegt werden.

Die Aggregationsrechnung erfolgt als separater Prozess über eine Warteschlange, damit keine konkurrierenden Aggregationsrechnungen auftreten können. Dies ist unbedingt erforderlich, da der erste Bearbeitungsschritt der Aggregationsfunktion das Löschen aller bereits zu der festgelegten Messwertmenge existierenden, aggregierten Werte ist. Überlappen diese Messwertmengen, könnte es sonst zum Deadlock, also einer Blockade der Datenbank, kommen (Elmasri & Navathe, 2016).

7.12 Reporterstellung EUA/Nitrat-RL

Da im Umweltkompartiment Fließgewässer mitunter kontinuierliche Messungen (z. B. der Temperatur) vorliegen, die EUA jedoch nur einen Wert pro Tag und Matrix akzeptiert, besteht der Wunsch, eine „validiert“-Flag für Messwerte mitzuführen. Mithilfe dieses Flags sollen bei der Reporterstellung nur diejenigen Messwerte aus kontinuierlichen Messungen verwendet werden, die validiert = True bzw. Wahr sind. Da analoge Konzepte bereits in der Vergangenheit im Umweltkompartiment Boden umgesetzt wurden und die Möglichkeit, Messwerte mit Validierungsergebnissen zu versehen sinnvoll ist, wurde dies für die Messwerte aller Kompartimente umgesetzt.

Es werden zwei validiert-Flags eingesetzt:

- ▶ Validität aus automatisierter QS (validity_automated)
- ▶ Validität aus Experteneinschätzung (validity_expert).

Generell wird der Validität aus Experteneinschätzung gegenüber der automatisiert erzeugten Validität der Vorrang eingeräumt, wenn Algorithmen zur Auswahl von Messwerten programmiert werden. Findet keine automatisierte QS statt, wie im Fall der Fließgewässer, wird das Flag nicht gesetzt. Das Flag für die Validität aus Experteneinschätzung wird generell nicht automatisiert gesetzt. Es kann über die Anwendungsoberfläche vom Nutzer gesetzt oder zurückgesetzt werden.

Derzeit sind im Reporting folgende Berichtspflichten umgesetzt:

- ▶ Nitratrictlinie
- ▶ WISE6 Disaggregated Data
- ▶ EU-Watch-List

Alle bislang implementierten Berichte arbeiten auf einem Zeitraum, für den der Bericht erstellt werden soll und auf den Messstellen eines auswählbaren Messnetzes. Die Berichte zu WISE6 Disaggregated Data und zur EU-Watch-List ermöglichen die Angabe eines Satzes von Parametern, der durch Hinzufügen einzelner Parameter oder von vorkonfigurierbaren Parametergruppen zusammengestellt werden kann.

Die Berichterstellung erfolgt asynchron in einer Warteschlange. Je nach Art des Berichtsmoduls werden Tabellen im .xlsx-Format oder XML-Dateien erstellt.

Eine Besonderheit stellt der Bericht zur EU-Watch-List dar. Er muss einen anderen Watch-List-Bericht und auch einen WISE6-Bericht zusätzlich integrieren können, darf jedoch keinerlei Duplikate bei den Messwerten enthalten, da solche zur Ablehnung beim Import in das Reportnet der EUA führen würden. Eine inkrementelle Datenlieferung, die das Problem verhindern würde, sieht Reportnet hier nicht vor.

8 Rollen und Rechte WaBo-Intranet

Für die Regelung des Zugriffs auf Teile der Anwendung und insbesondere, welche Operationen auf welchen Teilen und Daten der Anwendung ausgeführt werden dürfen, bedarf es der Festlegung von Rollen und damit verbundenen Rechten, sowie der Umsetzung dieser Festlegungen in der eigentlichen Anwendung.

8.1 Rollen von Benutzer:innen

Folgende Rollen können Benutzer:innen zugewiesen werden. Einige Rollen werden Benutzer:innen global zugewiesen, andere gelten nur in bestimmten Kontexten, z. B. kann ein:e Benutzer:in Verwalter:in für ein bestimmtes Umweltkompartiment sein, aber nicht für ein anderes.

Tabelle 1: Rollen von Benutzer:innen WaBo-Intranet

Rollenname	Kontext
Administrator:in	
Verwalter:in	Umweltkompartiment

8.2 Benutzer:innen-Rechte

Für folgende Operationen werden die Rechte von Benutzer:innen individuell bestimmt. Einige Rechte werden Benutzer:innen global zugewiesen, andere gelten nur in bestimmten Kontexten. Zu jeder Operation gibt es Bedingungen unter welchen die Benutzer:in die Rechte für die Operation hat. Grundsätzlich müssen Benutzer:innen authentifiziert (eingeloggt) sein, um überhaupt Operationen ausführen zu können. Operationen, die nicht explizit aufgelistet sind, sind allen Benutzer:innen **nicht erlaubt**.

Tabelle 2: Benutzer:innen-Rechte WaBo-Intranet

Operation	Kontext	Bedingung
Benutzer:in erstellen		Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment
Benutzer:in löschen	zu löschende:r Benutzer:in	Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment und die:der zu löschende:r Benutzer:in ist Mitglied keines Kompartiments.
Benutzer:in einsehen	einzusehende:r Benutzer:in	Benutzer:in ist Administrator:in oder Benutzer:in ist einzusehende:r Benutzer:in selbst.
Benutzer:innen auflisten	Benutzer:in, Umweltkompartiment	Benutzer:in ist Administrator:in, Verwalter:in für mindestens ein Umweltkompartiment oder Mitglied im selben Kompartiment.

Operation	Kontext	Bedingung
Benutzer:in Email oder Name bearbeiten	zu bearbeitende:r Benutzer:in	Benutzer:in ist Administrator:in oder Benutzer:in ist zu bearbeitende:r Benutzer:in selbst.
Benutzer:in die Administrator:in-Rolle zuweisen		Benutzer:in ist Administrator:in
Benutzer:in einem Kompartiment zuweisen	zu bearbeitende:r Benutzer:in, Kompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für betroffenes Umweltkompartiment.
Benutzer:in aus einem Kompartiment entfernen	zu bearbeitende:r Benutzer:in, Kompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für betroffenes Umweltkompartiment und zu bearbeitende:r Benutzer:in ist nicht Verwalter:in für das betroffene Kompartiment.
Benutzer:in die Verwalter:in-Rolle zuweisen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für das betroffene Umweltkompartiment
Benutzer:in die Verwalter:in-Rolle entziehen	Umweltkompartiment	Benutzer:in ist Administrator:in
Parameter erstellen oder ändern	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment
Parameter auflisten oder einsehen		
Parameteralias erstellen, auflisten, einsehen, ändern oder löschen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Mitglied in betroffenem Umweltkompartiment
Parametergruppe erstellen, auflisten, einsehen, ändern oder löschen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Mitglied in betroffenem Umweltkompartiment
Physikalische Einheit erstellen oder ändern		Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment
Messgröße erstellen		Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment
Messgröße auflisten oder einsehen		
Liefervorgänge auflisten oder einsehen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Mitglied in betroffenem Umweltkompartiment
Feedback erstellen, auflisten oder einsehen		

Operation	Kontext	Bedingung
Feedback ändern oder löschen	Umweltkompartiment, Benutzer:in	Benutzer:in ist Administrator:in oder Verwalter:in mindestens eines Umweltkompartiments in dem die:der Benutzer:in, für welche:n das Feedback ist, Mitglied ist
Umweltkompartimente auflisten oder einsehen		
Bundesländer auflisten oder einsehen		
Matrizen auflisten oder einsehen		
Messnetz erstellen, auflisten, einsehen, ändern oder löschen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Mitglied in betroffenem Umweltkompartiment
Messstelle erstellen, auflisten, einsehen, ändern oder löschen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Mitglied in betroffenem Umweltkompartiment
Aufnahmesituation erstellen, auflisten, einsehen, ändern, löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Horizonteigenschaften - Ausgangsgesteinsanteil erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Horizonteigenschaften - Gesteinskennzeichnung erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Horizonteigenschaften - Kornfraktion erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Profilkennzeichnung erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Substratsystematik - Klasse erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Substratsystematik - Typ erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Substratsystematik - Subtyp erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden

Operation	Kontext	Bedingung
Bodenhorizont erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
Bodenhorizonteigenschaft erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Boden
See erstellen, auflisten, einsehen, ändern oder löschen		Benutzer:in ist Administrator:in oder Mitglied im Umweltkompartiment Seen
Ordner in Datawarehouse erstellen		Benutzer:in ist Administrator:in oder Mitglied in mindestens einem Umweltkompartiment
Ordner in Datawarehouse auflisten und einsehen		Alle Benutzer:innen die Zugang zum Intranet haben
Ordner in Datawarehouse ändern		Unlöschar Kennzeichnung darf nicht entfernt werden, Benutzer:in ist Administrator:in, Verwalter:in oder Ersteller:in des Ordners
Ordner in Datawarehouse entfernen		Ordner darf nicht unlöschar sein, Benutzer:in muss über die Rechte verfügen alle enthaltenen Ordner und Dateien zu löschen
Datei in Datawarehouse erstellen		Benutzer:in ist Administrator:in oder Mitglied in mindestens einem Umweltkompartiment
Datei in Datawarehouse auflisten und einsehen		Alle Benutzer:innen die Zugang zum Intranet haben
Datei in Datawarehouse ändern oder entfernen		Benutzer:in ist Administrator:in, Ersteller:in der Datei oder Verwalter:in eines Kompartiments, welchem die Datei zugeordnet ist

9 Rollen und Rechte WaBo-DMZ

Selbstverständlich bedarf es auch im extern, v.a. für Datenlieferanten verfügbaren Teil von FIS Wasser-Boden der Regelung des Zugriffs auf Teile der Anwendung und die für den Anwender erlaubten Operationen.

9.1 Rollen von Benutzer:innen

Folgende Rollen können Benutzer:innen zugewiesen werden. Einige Rollen werden Benutzer:innen global zugewiesen, andere gelten nur in bestimmten Kontexten, z. B. kann ein:e Benutzer:in Verwalter:in für ein bestimmtes Umweltkompartiment sein, aber nicht für ein anderes.

Tabelle 3: Rollen von Benutzer:innen WaBo-DMZ

Rollenname	Kontext
Administrator:in	
Verwalter:in	Umweltkompartiment
Mitglied	Umweltkompartiment

9.2 Rollen von Organisationen

Benutzer:innen können einer oder mehreren Organisationen zugewiesen werden. Organisationen sind alle Entitäten, die über WaBo-DMZ Berichte einreichen können, z. B. Landesumweltämter.

Analog zu Benutzer:innen können auch Organisationen Rollen kontextbezogen zugewiesen werden. Folgende Rollen sind vorgesehen:

Tabelle 4: Rollen von Organisationen

Rollenname	Kontext
Berichtertatter:in	Umweltkompartiment, Berichtsart, Bundesländer, Messnetze

9.3 Benutzer:innen-Rechte

Für folgende Operationen werden die Rechte von Benutzer:innen individuell bestimmt. Einige Rechte werden Benutzer:innen global zugewiesen, andere gelten nur in bestimmten Kontexten. Zu jeder Operation wird festgelegt unter welcher Bedingung Benutzer:innen die Operation ausführen dürfen. Grundsätzlich dürfen nur angemeldete Benutzer:innen Operationen ausführen. Operationen, die nicht explizit aufgelistet sind, sind allen Benutzer:innen **nicht erlaubt**.

Tabelle 5: Benutzer:innen-Rechte WaBo-DMZ

Operation	Kontext	Bedingung
Benutzer:in erstellen		Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment

Operation	Kontext	Bedingung
Benutzer:in auflisten	Umweltkompartiment	Benutzer:in ist Administrator:in, Verwalter:in für mindestens ein Umweltkompartiment oder Benutzer:in teilt gemeinsame(s) Umweltkompartiment(e)
Benutzer:in einsehen	Umweltkompartiment	Benutzer:in ist Administrator:in, Verwalter:in für betroffenes Umweltkompartiment oder Benutzer:in sieht eigenes Profil ein
Benutzer:in löschen	zu bearbeitende:r Benutzer:in	Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment und die:der zu bearbeitende/löschende Benutzer:in hat keine Rollenzuweisung und ist Mitglied keiner Organisation.
Benutzer:in einer Organisation zuweisen oder aus einer Organisation entfernen	zu bearbeitende:r Benutzer:in, Organisation	Benutzer:in ist Administrator:in oder Verwalter:in für alle Umweltkompartimente, für die die Organisation Rollenzuweisungen besitzt.
Benutzer:in die Administrator:in-Rolle zuweisen		Benutzer:in ist Administrator:in
Benutzer:in die Verwalter:in-Rolle zuweisen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für das betroffene Umweltkompartiment
Benutzer:in die Verwalter:in-Rolle entziehen		Benutzer:in ist Administrator:in
Email und Namen von Benutzer:in ändern		Benutzer:in ist Administrator oder Benutzer:in bearbeitet eigenes Profil
Organisation erstellen		Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment
Organisation auflisten oder einsehen	aufzulistende Organisation	Benutzer:in ist Administrator:in, Verwalter:in für mindestens ein Umweltkompartiment oder Benutzer:in ist Mitglied in Organisation
Organisation löschen	zu bearbeitende Organisation	Benutzer:in ist Administrator:in oder Verwalter:in für mindestens ein Umweltkompartiment und die zu bearbeitende Organisation hat keine Mitglieder und keine Lieferberechtigungen.

Operation	Kontext	Bedingung
Lieferberechtigung erstellen	Umweltkompartiment, Berichtsart, Bundesländer, Messnetze	Benutzer:in ist Administrator:in oder Verwalter:in für das betroffene Umweltkompartiment
Lieferberechtigung auflisten oder einsehen	Umweltkompartiment, Organisation	Benutzer:in ist Administrator:in, Verwalter:in von mindestens einem Umweltkompartiment oder Mitglied in betroffener Organisation
Lieferberechtigung löschen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für das betroffene Umweltkompartiment
Berichtskampagne erstellen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für das betroffene Umweltkompartiment, für das eine Berichtskampagne erstellt werden soll.
Berichtskampagne auflisten oder einsehen	Umweltkompartiment, Lieferberechtigung, Organisation	Benutzer:in ist Administrator:in, Verwalter:in für das betroffene Umweltkompartiment oder Mitglied in Organisation für die Lieferberechtigung zu Kampagne besteht.
Berichtskampagne bearbeiten und löschen	zu bearbeitende Berichtskampagne	Benutzer:in ist Verwalter:in für das betroffene Umweltkompartiment, für das Umweltkompartiment der zu verwaltenden Berichtskampagne.
Neuen Berichtsvorgang erstellen	Umweltkompartiment, Berichtsart, Bundesländer, Messnetze	Benutzer:in ist Administrator:in, Verwalter:in für betroffenes Umweltkompartiment oder Mitglied in Organisation mit entsprechender Lieferberechtigung hat die Rolle "Berichterstatter:in" in diesem Kontext
Berichtsvorgang einsehen (beinhaltet sämtliche Metadaten, die Importdatei, alle Prüfprotokolle)	Importvorgang	Benutzer:in ist Administrator:in, Verwalter:in für betroffenes Umweltkompartiment oder Mitglied in Organisation, die diesen Berichtsvorgang erstellt hat.
Berichtspflicht erstellen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für betroffenes Umweltkompartiment
Berichtspflicht auflisten oder einsehen	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für betroffenes Umweltkompartiment

Operation	Kontext	Bedingung
Berichtspflicht ändern	Umweltkompartiment	Benutzer:in ist Administrator:in oder Verwalter:in für betroffenes Umweltkompartiment
Umweltkompartiment erstellen		Benutzer:in ist Administrator:in

10 Datenübertragung und Modellkonsolidierung

Die Entwicklung der automatisierten Datenübertragung der Altdaten aus ihren jeweiligen Datenbanken in die Datenbank des FIS-WaBo-Intranet ist der Zeitpunkt, an dem die wesentlichen Daten aus den alten Strukturen in die neue Struktur überführt werden. Im Zweifelsfall zeigt sich spätestens hier schlechte Modellierung der Altdaten, wie z. B. Listen mit bedeutungsübergreifend gemischten Einträgen.

Bei der Modellierung von FIS-WaBo wurde darauf geachtet, dass solche Listen möglichst vermieden werden. Dies ließ sich jedoch aufgrund der Anforderungen nicht vollständig durchhalten, die explizit die Konservierung von teilweise bereits in den zugrundeliegenden Gesetzen schlecht beschriebenen Vorgehensweisen forderten.

Hier liegt die **wesentliche Erkenntnis dieses Forschungsvorhabens**:

Bei der Implementierung von Prozessen und technischen Werkzeugen, die aus den Inhalten von Gesetzen oder anderen Rechtserkenntnisquellen resultieren, ist nicht den mitunter technisch unsauberen Formulierungen in diesen Gesetzen zu folgen, sondern auf eine technische saubere und für die Anwender verständliche Umsetzung zu achten, die die rechtlichen Vorgaben erfüllt. Dazu sollte also, falls sinnvoll, von den Buchstaben der Vorlage abgewichen und eine saubere Methodik gewählt werden. Das nachfolgende Beispiel zeigt dies anhand eines eins zu eins aus dem Gesetzestext abgeleiteten Modells, das Potenzial zur Verbesserung aufzeigt.

10.1 Modellpotenzial

Die Messwerte eines der in das FIS Wasser-Boden überführten Umweltkompartimente kennt in der Urprungsanwendung folgende Listen-Attribute, die bzgl. ihrer Benennung und ihrer Inhalte optimiert werden können:

► Messart

- E = Einzelprobe
- M = Mischprobe
- K = Kontinuierliche Messung

Die Listenelemente sind verständlich und thematisch homogen. (Der Eintrag „Kontinuierliche Messung“ könnte noch in „Kontinuierliche Probenahme“ umbenannt werden. Das ist jedoch eher kosmetischer Natur.) Die Liste ist folglich sauber modelliert. Jedoch ist der Attributname „Messart“ ohne die Listeninhalte nicht verständlich. In anderen Umweltkompartimenten ist der Name „Probenahmeart“ für die Differenzierung von Einzel- und Mischproben etc. geläufig. Da sich die Liste auch auf die Art der Probenahme bezieht, wäre der Begriff „Probenahmeart“ zielführender.

► Probeart

- B = Biota-Fischfilet
- F = Filterrückstand
- S = Absetzbecken
- W = Wasserphase
- Z = Zentrifuge

Der Name des Attributs „Probeart“ ist bei der Erlangung der Erkenntnis, was hier beschrieben wird, nicht hilfreich. Die Liste ist inhomogen, da sie sich einerseits auf den Probengegenstand (einen Fisch oder Wasser) bezieht, diese jeweils genauer unterteilt (Filet des Fisches, Wasserphase oder Filtrerrückstand) und dabei gleichzeitig die Art der Probenaufbereitung (filtriert, zentrifugiert oder abgesetzt) beschreibt. Dabei unterscheidet die Liste nicht zwischen filtrierter Wasserphase und Gesamtphase. Daher ist es auch schwer möglich, einen passenderen Namen für das Attribut zu benennen. Diese Liste ist ein Lehrbuchbeispiel dafür, wie Attribute nicht modelliert werden sollten.

► Parameter5

- 1 = Gesamtprobe
- 3 = abgesetzt, Überstand für Analyse
- 6 = gelöst
- 8 = Fische
- 9 = Muscheln
- 0 = Schwebstoff/Sediment

Der Name des Attributs gibt keinen Hinweis auf den Inhalt, der hier beschrieben werden soll. Der Begriff Parameter wird oft für eine Messgröße in Kombination mit ihrer Einheit und gelegentlich auch der Art der Probenaufbereitung oder der Analytik verwendet. Andererseits ist der Begriff Parameter auch synonym mit dem Begriff Attribut und steht für benannte Kriterien, mit deren Hilfe die Art eines Objekts beschrieben wird. Die Liste zeigt jedoch, dass es hier um das beprobte Objekt, man könnte auch sagen um den Gegenstand der Beprobung geht. Gelegentlich wird in anderen Umweltkompartimenten der Begriff Matrix dafür verwendet. Man könnte sicher auch Probenobjekt oder Probengegenstand als Namen verwenden und dem Betrachter so mehr Informationen an die Hand geben. Für den Eintrag „6 = gelöst“ könnte auch „6 = gelöst, Filtrat“ verwendet werden.

Der folgende Vorschlag zur Re-Modellierung von Parameter5 (Matrix) und Probenart (Probenaufbereitung) fand keine Zustimmung:

► Matrix

- 1 = Gesamtgehalt
- 6 = gelöste Konzentration
- 8 = Fisch
- 9 = Muscheln
- 0 = Schwebstoff/Sediment/Rückstand

► Probenaufbereitung

- filtriert
- zentrifugiert
- 2 h abgesetzt

- homogenisiert

Das Element „3 = abgesetzt, Überstand für Analyse“ geht in der gelösten Konzentration auf und bekommt die entsprechende Probenaufbereitung.

Der Vollständigkeit halber könnte für Fische und Muscheln auch noch die Probenaufbereitung „filetiert“ aufgenommen werden. Jedoch gibt es für Biota derzeit keine andere Aufbereitung, weshalb dies nicht zwingend erforderlich ist.

Die zahlreichen Probenaufbereitungen des Bodens umfassen mindestens auch „filtriert“ und „zentrifugiert“, so dass bereits ein gemeinsamer Teil zwischen allen hier betrachteten Umweltkompartimenten identifiziert ist. Wie in allen solchen Fällen gibt es eine Hauptliste, von der Umweltkompartiment-spezifische Listen abgeleitet werden, damit jede Nutzergruppe nur die Listenelemente sieht (bzw. ansehen muss), die im jeweiligen Umweltkompartiment gebraucht werden.

11 Datenmodelle

Das FIS Wasser-Boden wird aufgrund der Sicherheits- und Nutzungsanforderungen in die Anwendungen WaBo-DMZ und WaBo-Intranet unterteilt. Diese haben jeweils eigene Datenmodelle, die sich aufgrund der Notwendigkeit des Austauschs von Daten, z. B. zu Datenlieferungen, zum Teil überschneiden.

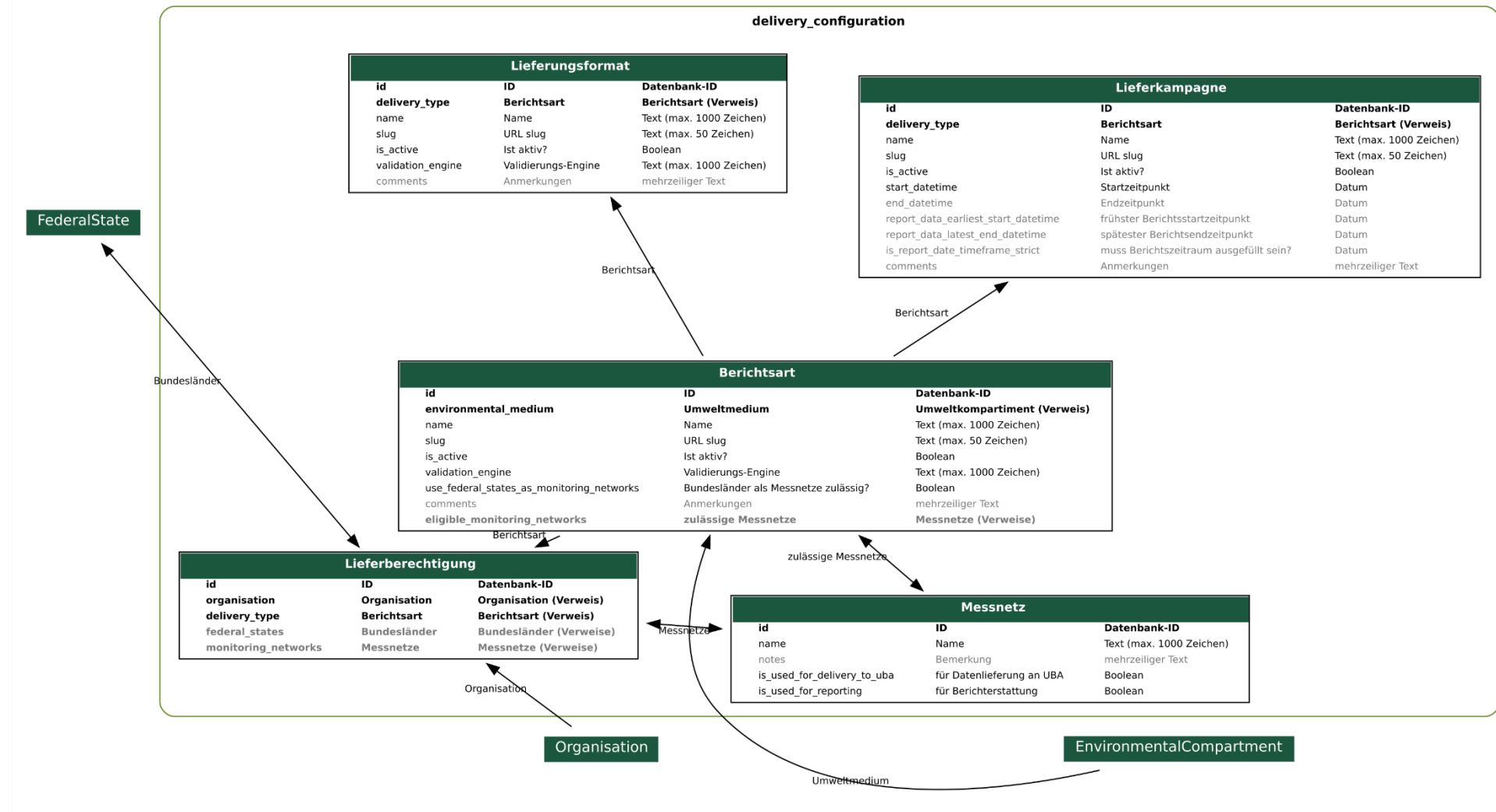
11.1 Datenmodell FIS Wasser-Boden DMZ (WaBo-DMZ)

WaBo-DMZ besitzt die vier Teildatenmodelle

- ▶ Datenlieferungskonfiguration (delivery_configuration)
- ▶ Datenlieferungsprozess (delivery_process)
- ▶ Listen (list) und
- ▶ Benutzer:innen (account)

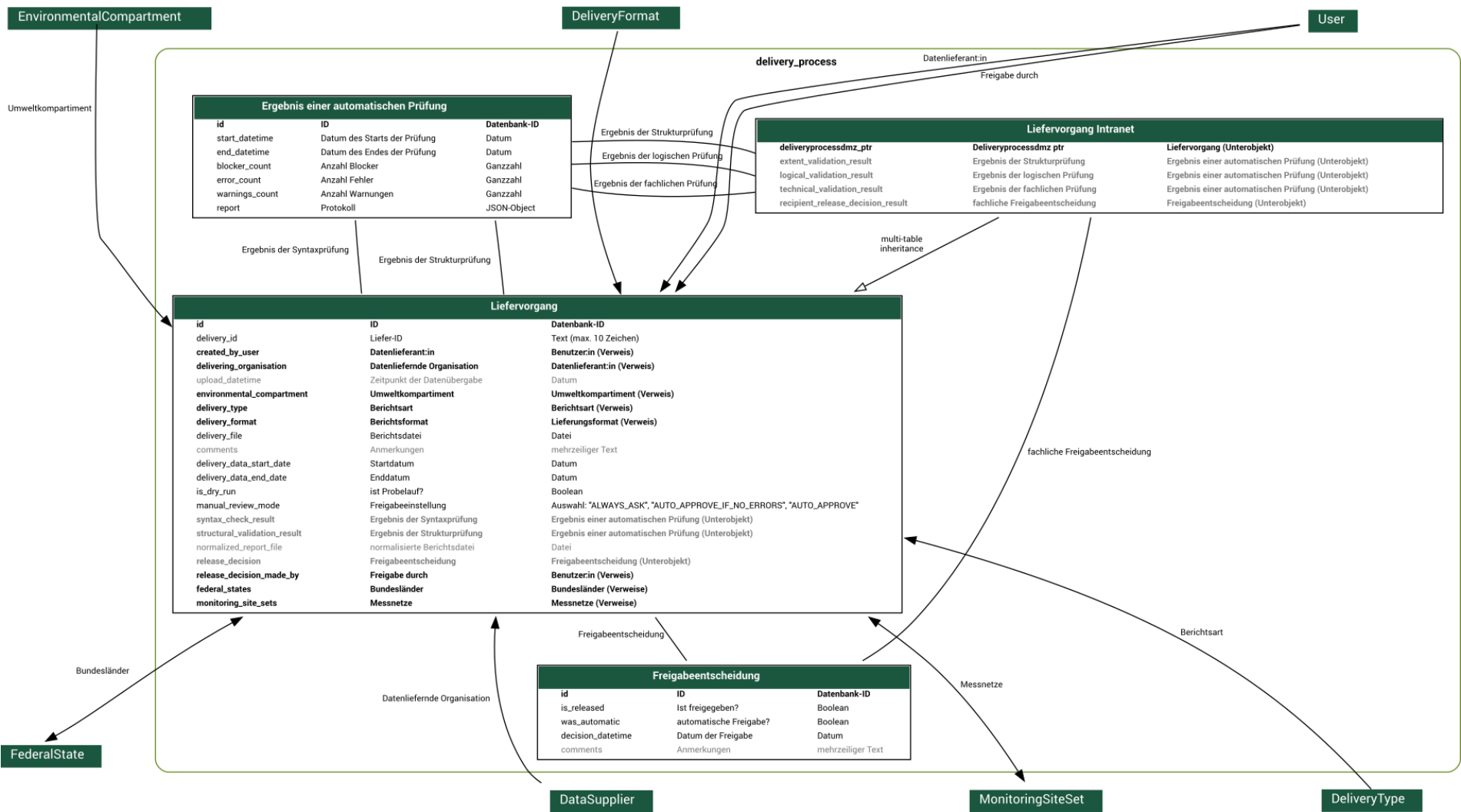
die wie folgt festgelegt werden:

Abbildung 7: Datenmodell Datenlieferungskonfiguration (delivery_configuration)



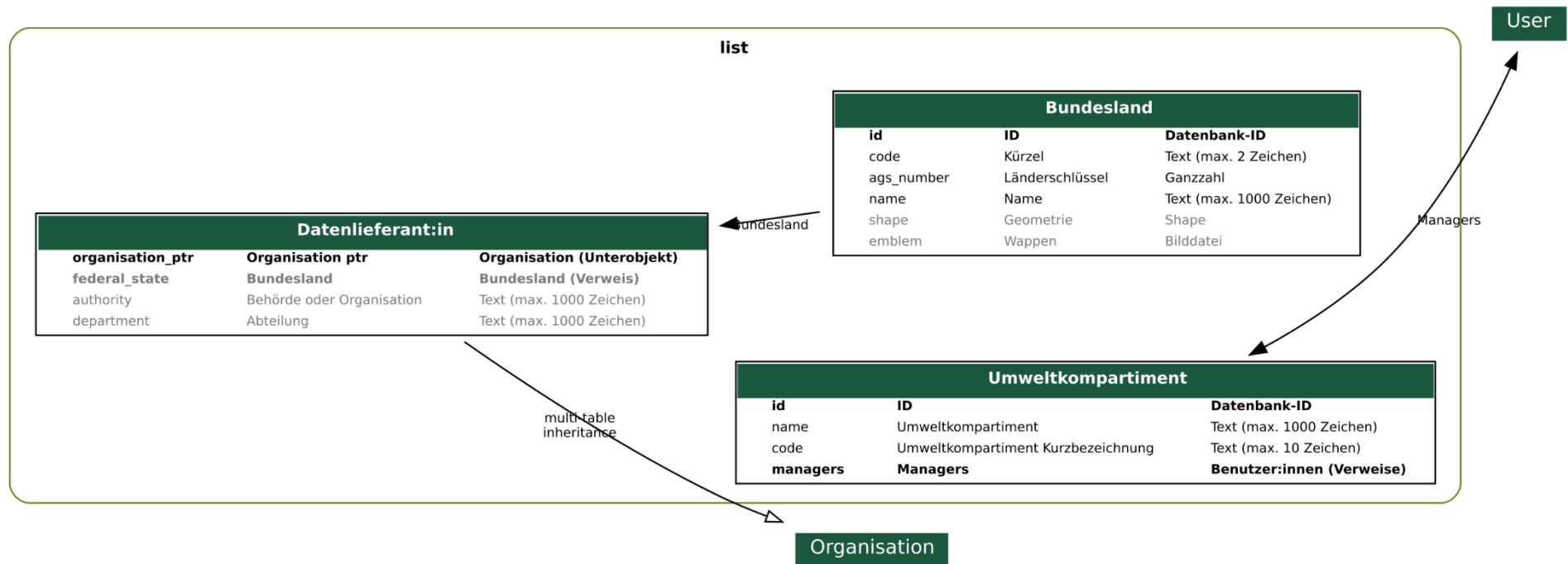
Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Abbildung 8: Datenmodell Datenlieferungsprozess (delivery_process)



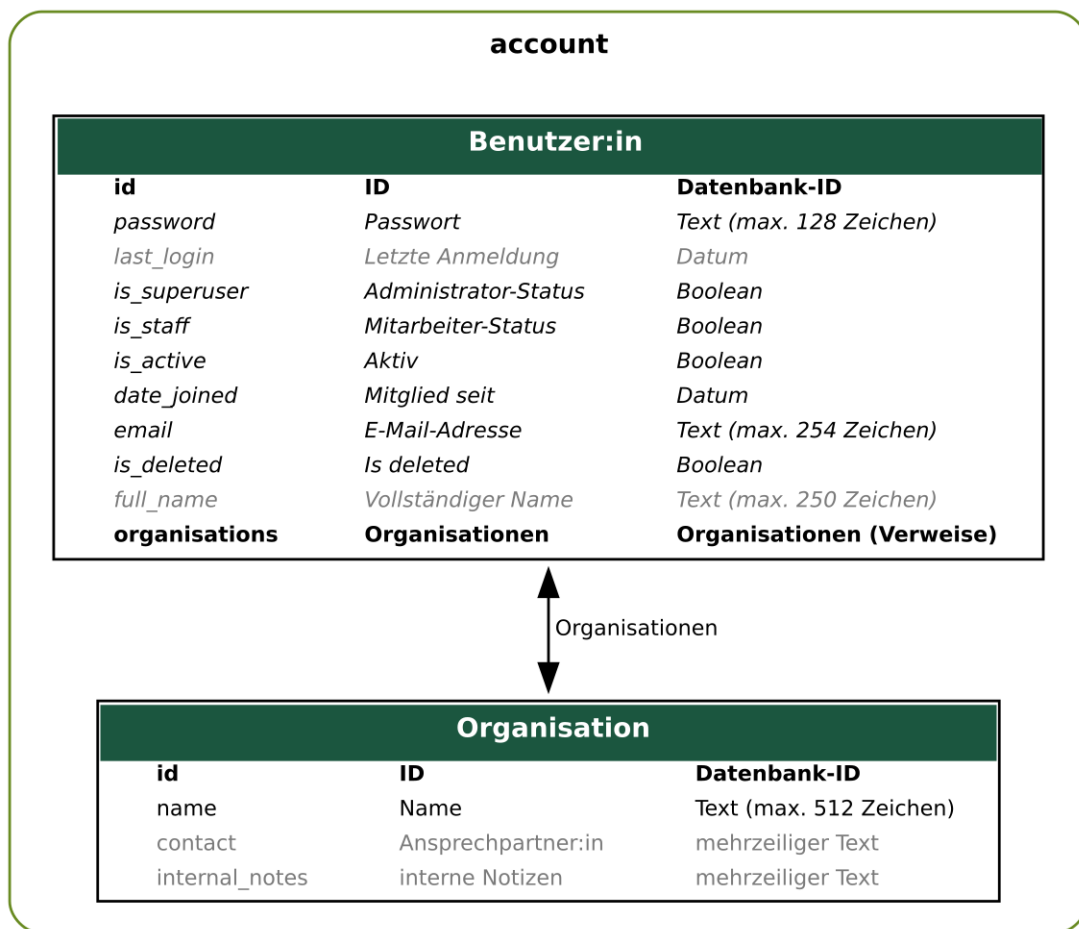
Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Abbildung 9: Datenmodell Listen (list)



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

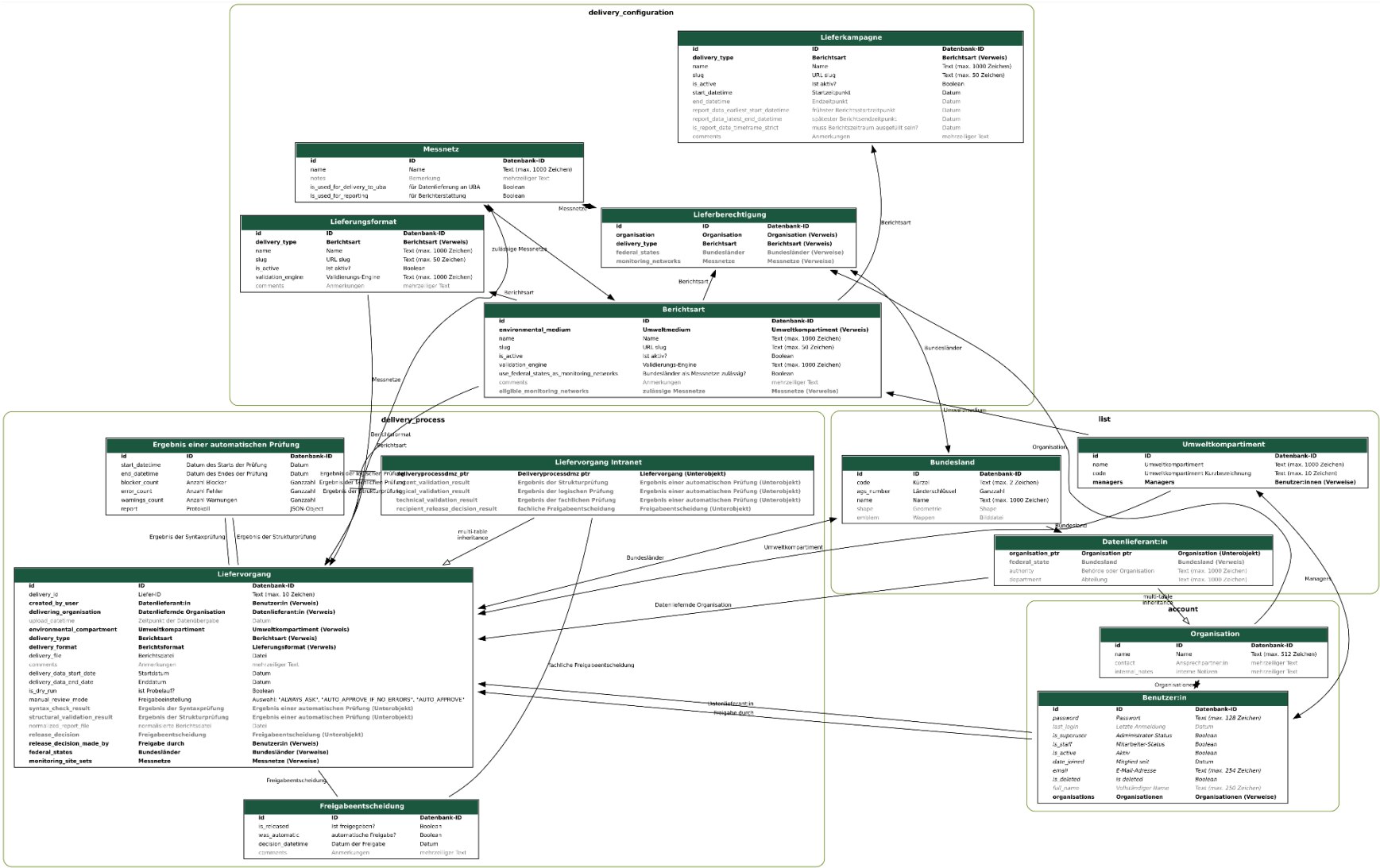
Abbildung 10: Datenmodell Benutzer:innen (account)



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Das Gesamtdatenmodell ist zur besseren Lesbarkeit auch im separaten ZIP-Container **Be-richtsgrafiken.zip** zum Abschlussbericht hinterlegt wie folgt festgelegt worden:

Abbildung 11: Gesamtdatenmodell WaBo-DMZ



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

11.2 Datenmodell FIS Wasser-Boden (WaBo-Intranet)

WaBo -Intranet besitzt die Teildatenmodelle

- ▶ Stammdaten (master_data)
- ▶ Parameter (measurand)
- ▶ Messungen (measurement)
- ▶ Metadaten (metadata)
- ▶ Datenlieferungskonfiguration (delivery_configuration)
- ▶ Listen (list) und
- ▶ Benutzer:innen (account).

Um einen Ausgangspunkt für die Detailmodellierung zu schaffen, wurde zunächst ein Kerndatenmodell entwickelt. Dieses enthält die wesentlichen Tabellen und deren Beziehungen untereinander:

Abbildung 12: Kerndatenmodell FIS Wasser-Boden Intranet

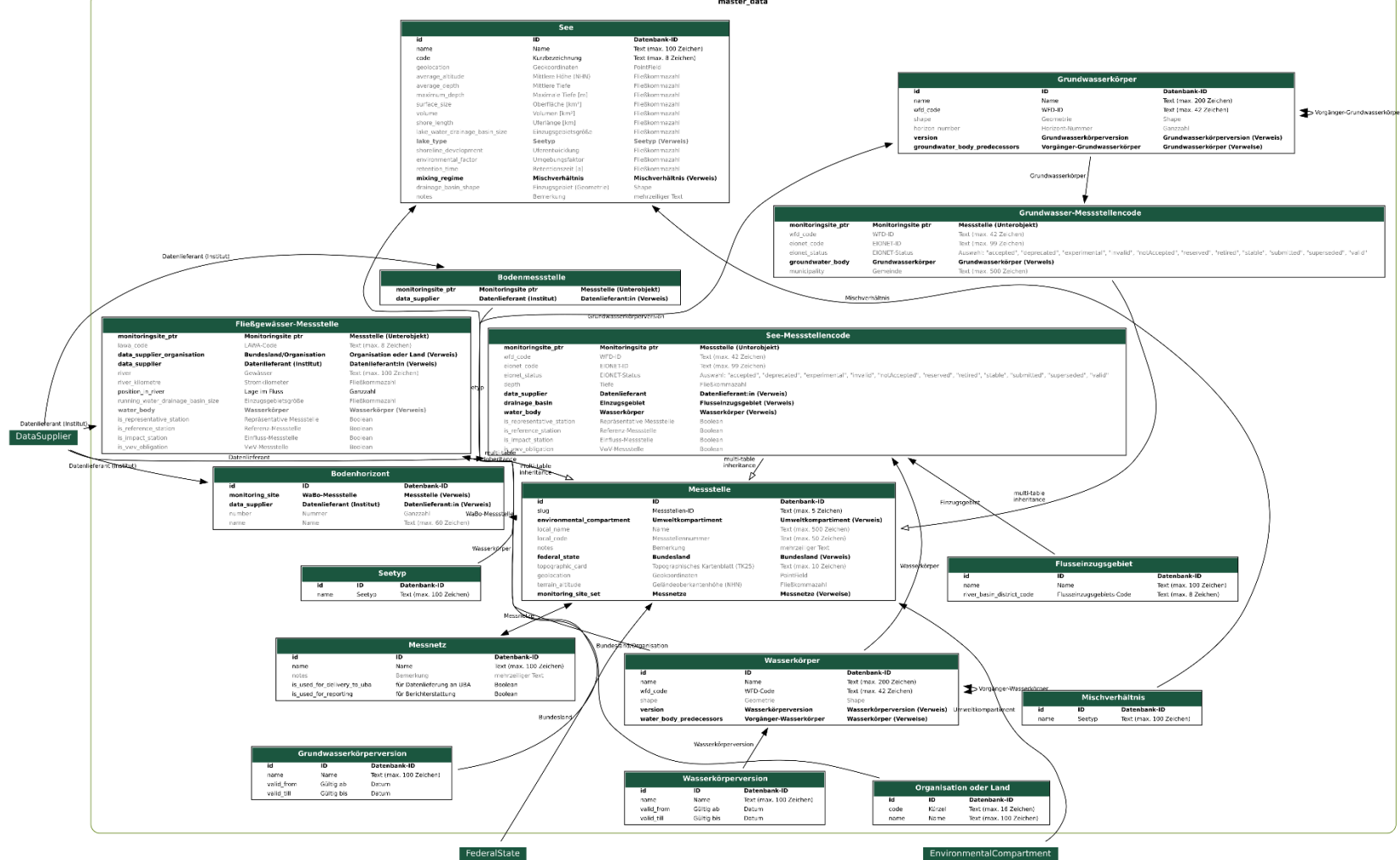


Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Aus dem Kerndatenmodell wurden die nachfolgend dargestellten Datenmodelle entwickelt und diejenigen Datenmodelle ergänzt, die in der Kerndatenmodellsicht ohne Bedeutung sind.

In den Datenmodellen ist speziell für die Grundwassermessstellen eine Versionierung im Sinne einer Historienverwaltung und auch eine Vorgänger-Nachfolger-Beziehung (0-1:0-1) vorgesehen, da dies für Grundwasser zwingend erforderlich ist¹³.

¹³ Insbesondere die korrekte Behandlung von Objekthistorien ist gewissenhaft zu berücksichtigen, wenn Abfragen und Reports keine unbeabsichtigten Messwertduplikate liefern sollen. Auch bei der Abfrage der Anzahl der Objekte (hier Grundwassermessstellen), können zu hohe Zahlen ermittelt werden, wenn der Zeitpunkt nicht berücksichtigt wird oder die Gültigkeitszeiträume eines Objekts sich überlappen.



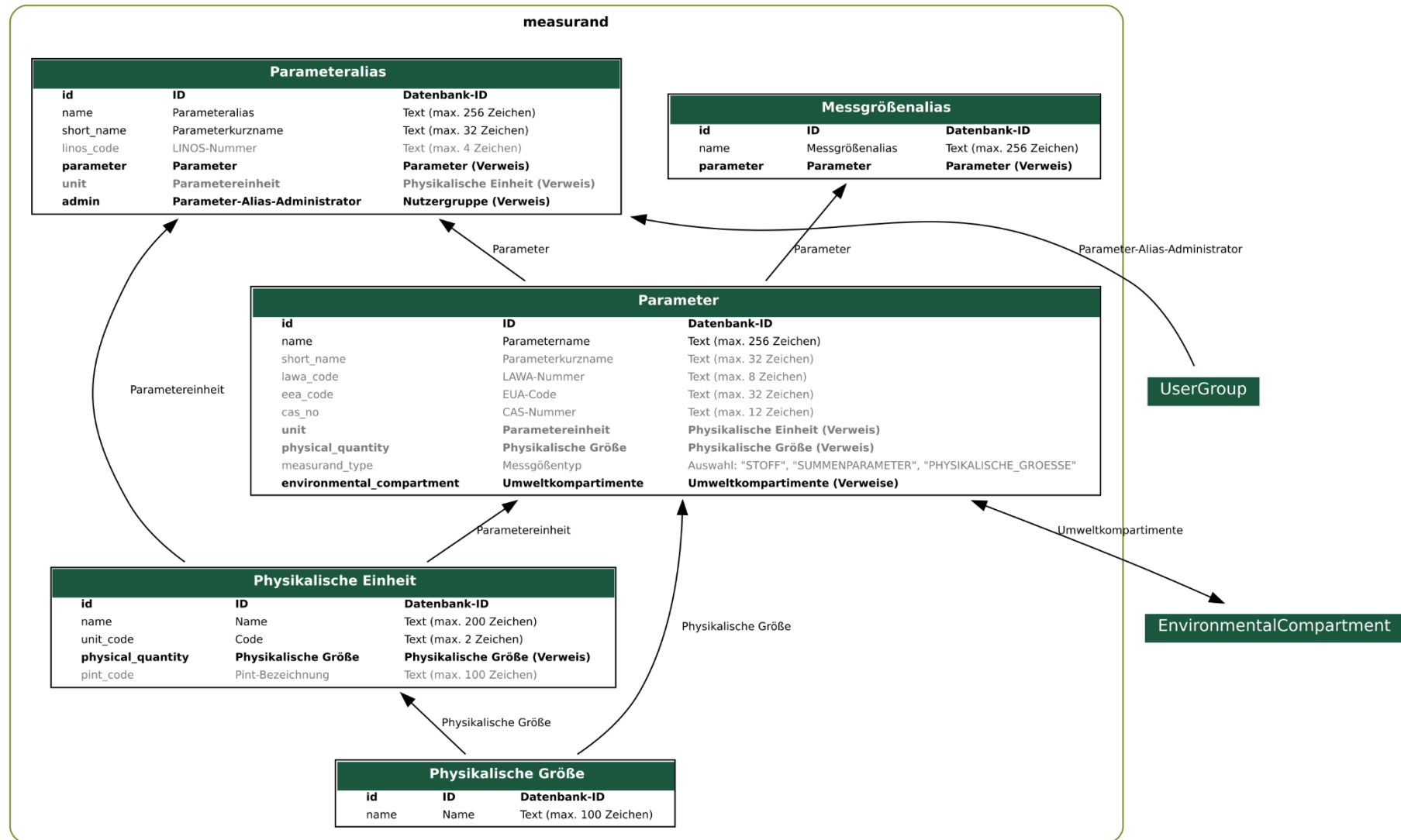
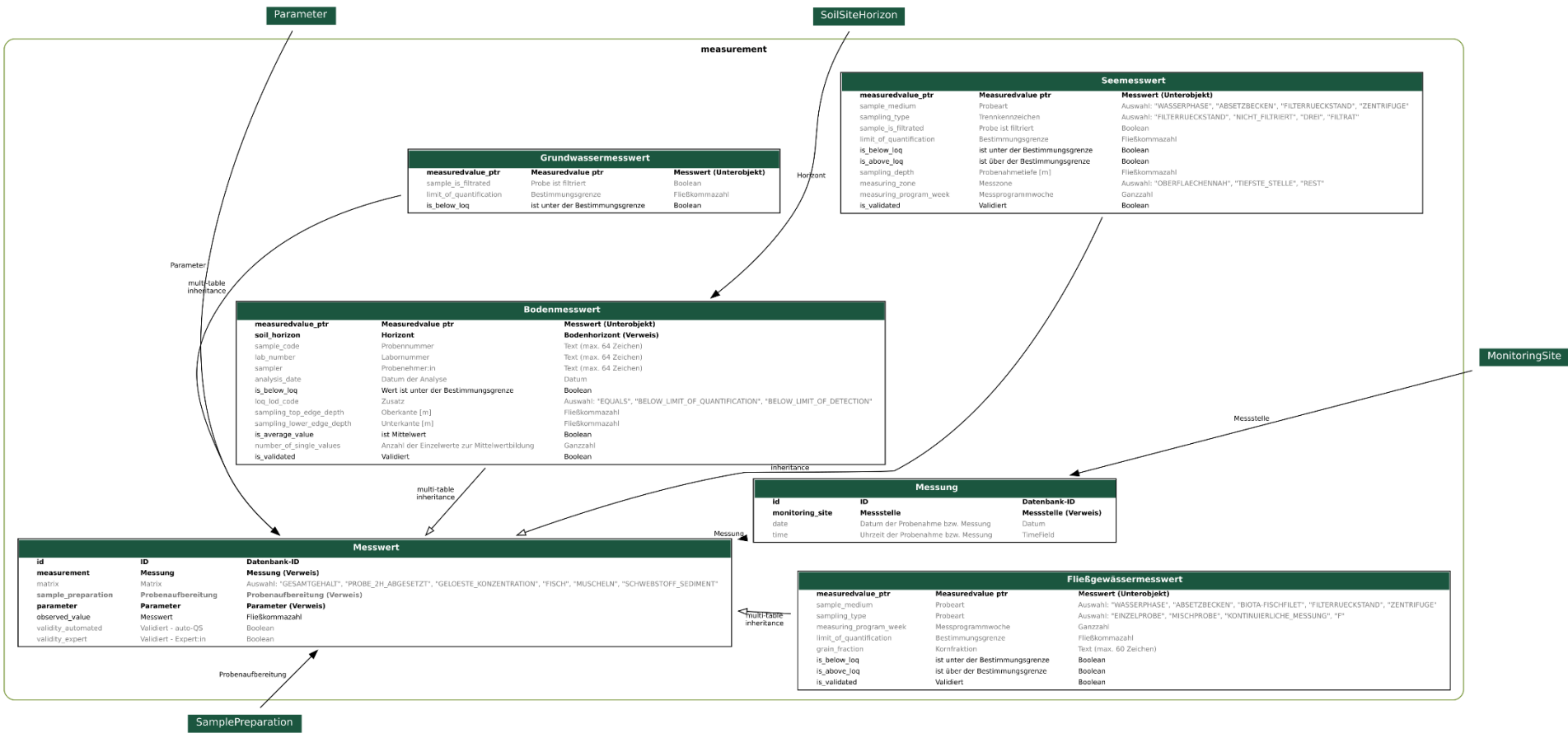
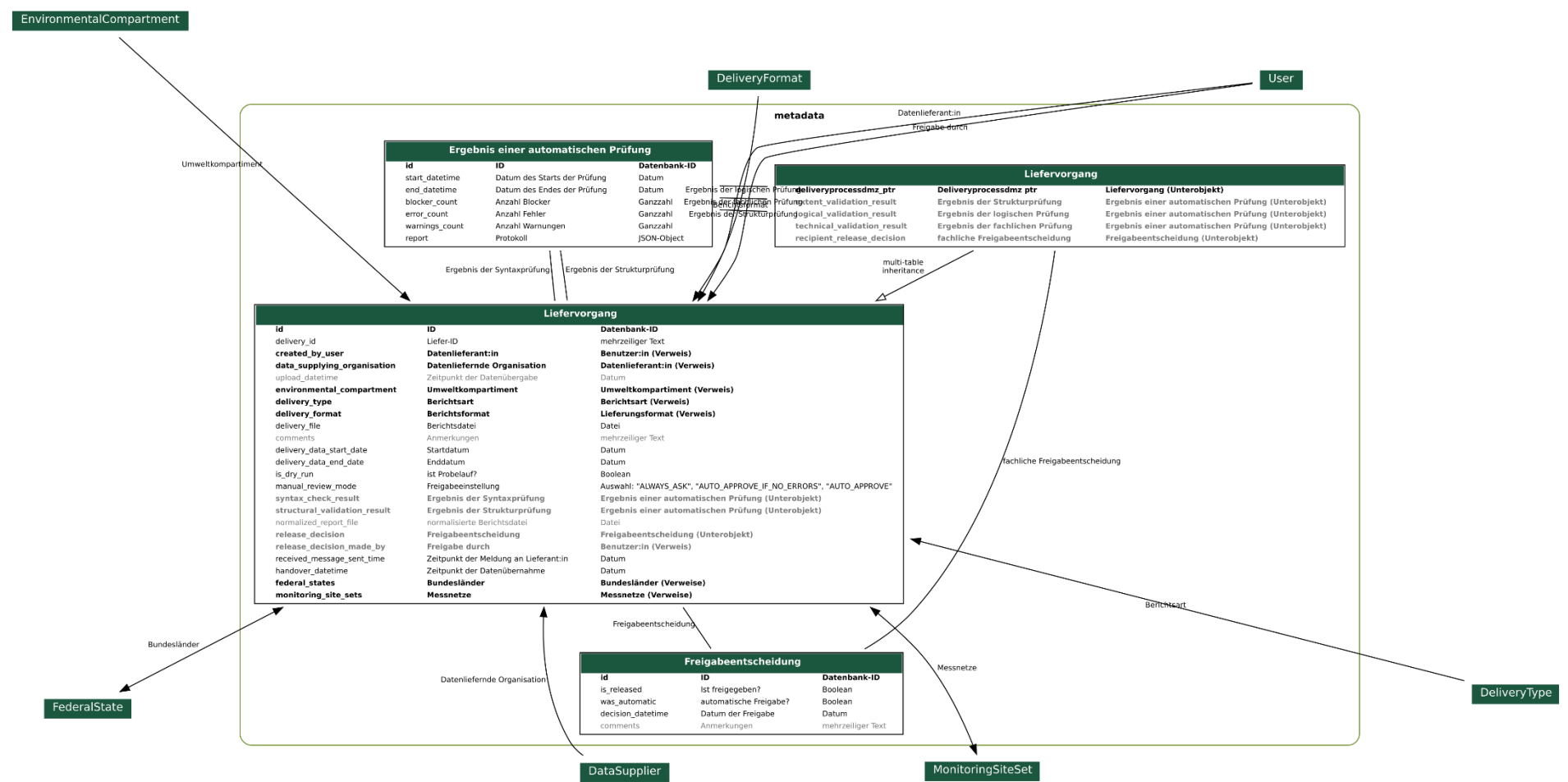


Abbildung 15: Datenmodell Messungen (measurement)



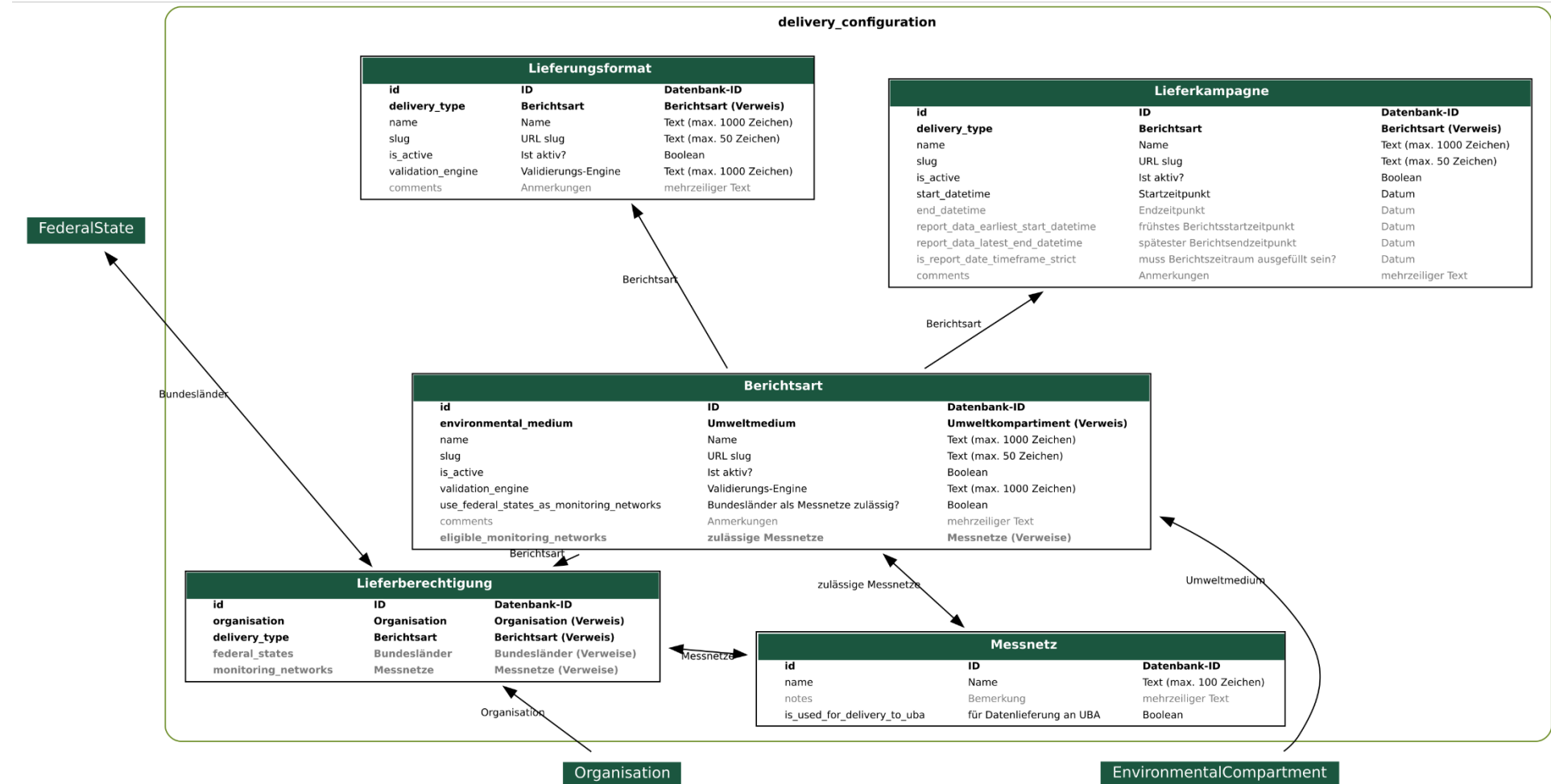
Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Abbildung 16: Datenmodell Metadaten (metadata)



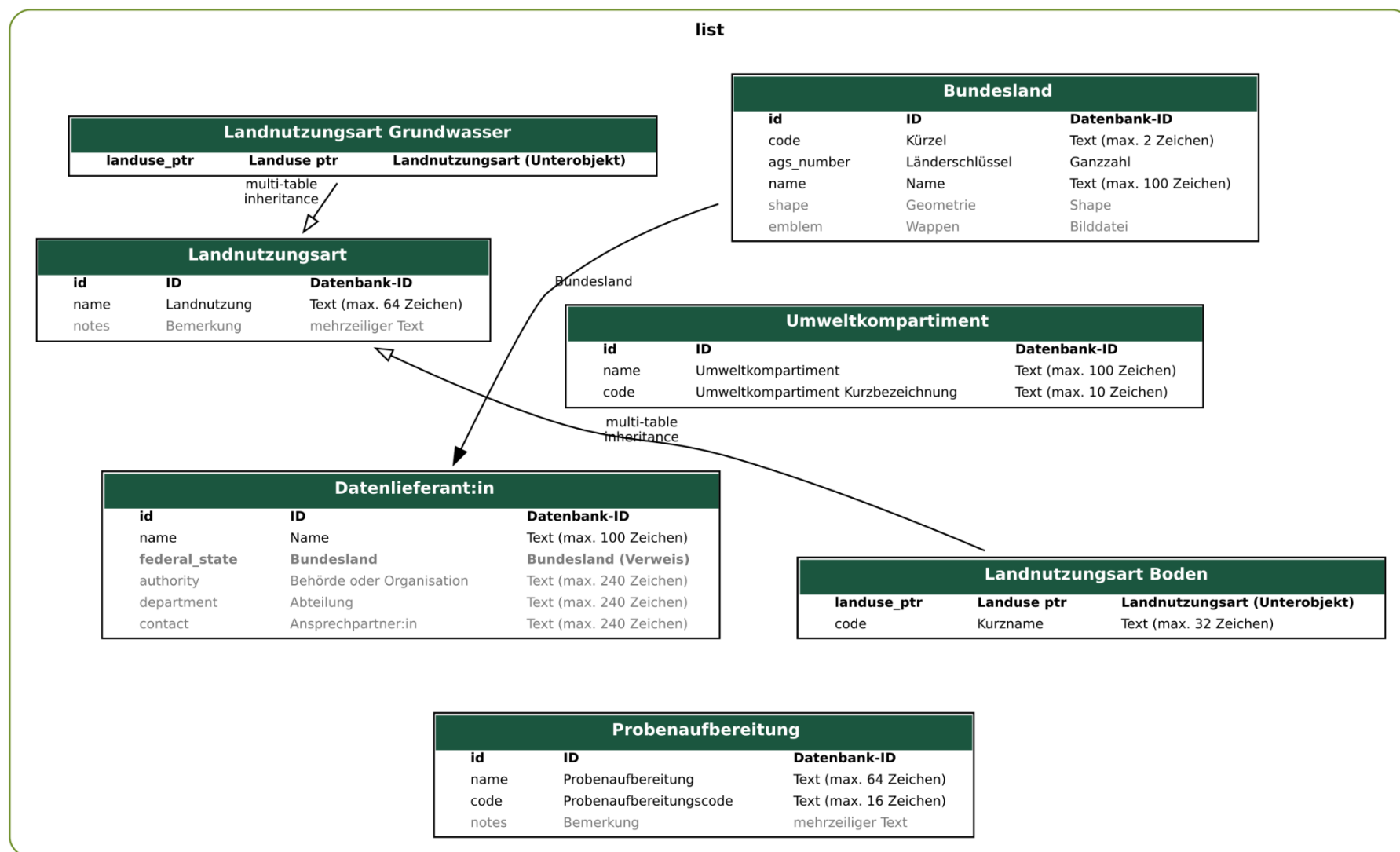
Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Abbildung 17: Datenmodell Datenlieferungskonfiguration (delivery_configuration)



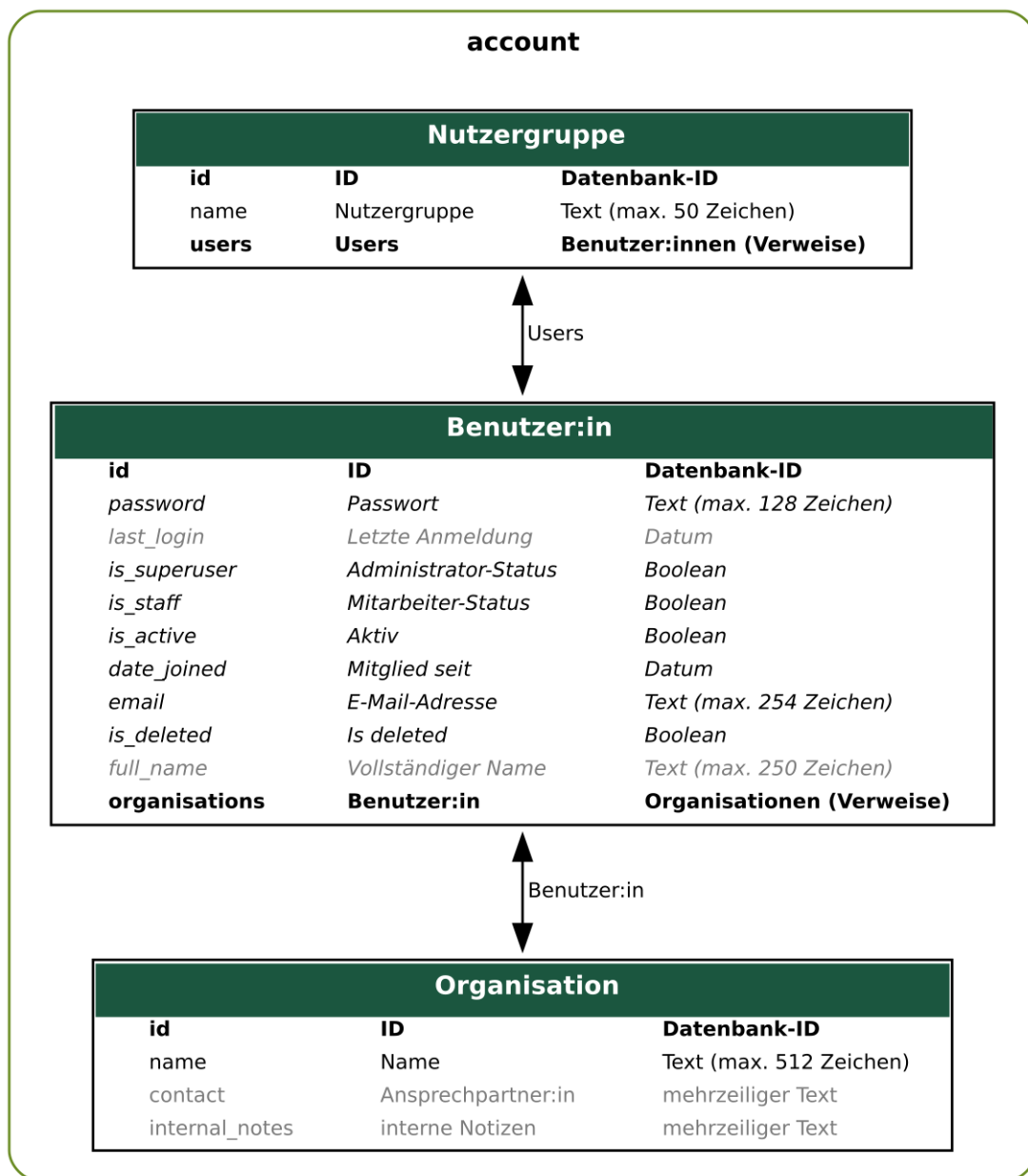
Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Abbildung 18 Datenmodell Listen (list)



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

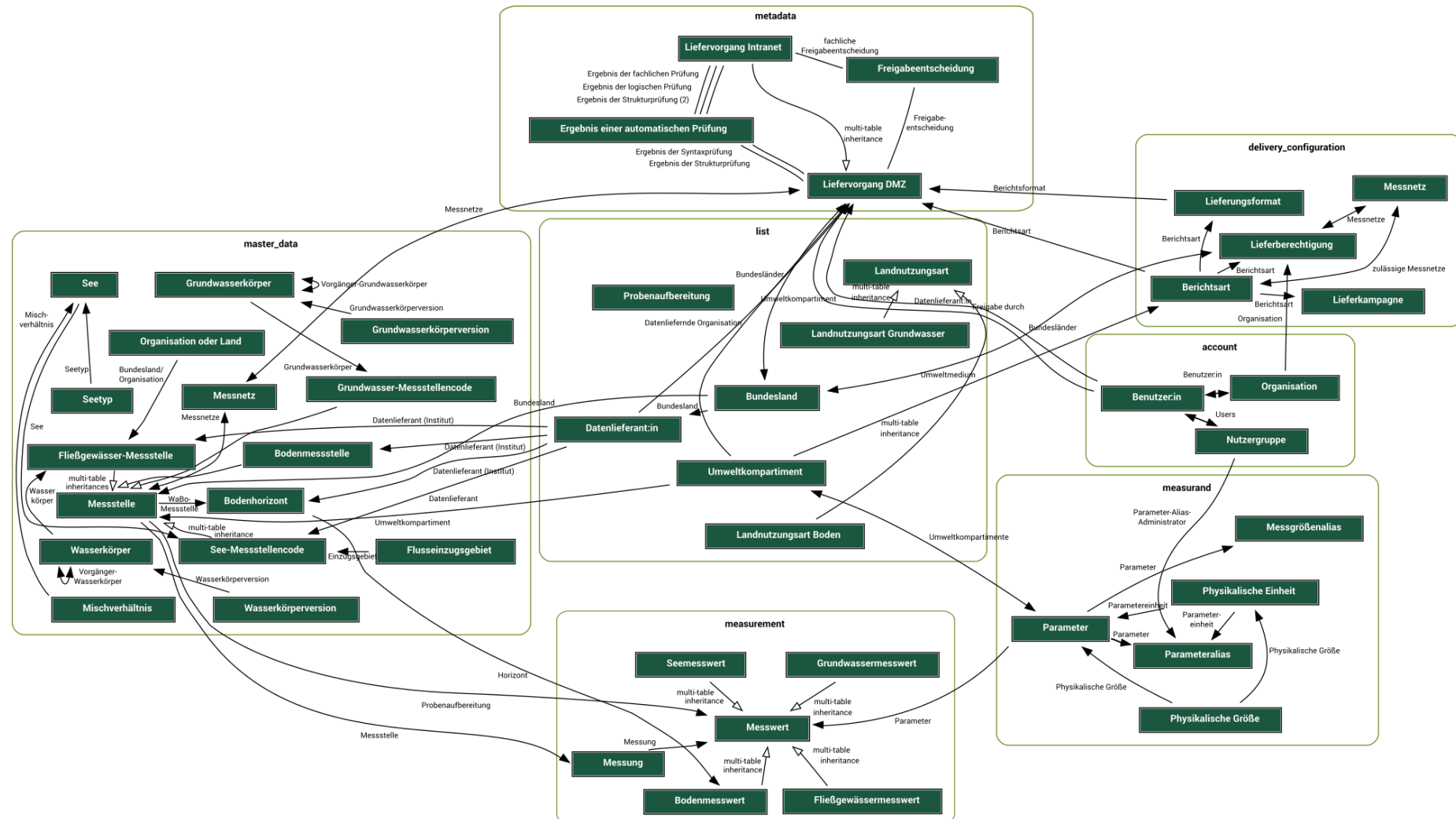
Abbildung 19: Datenmodell Benutzer:innen (account)



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Das Gesamtdatenmodell ist wie folgt festgelegt worden:

Abbildung 20: Gesamtdatenmodell WaBo-Intranet



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

11.2.1 Erläuterungen

Das Teildatenmodell Listen („list“) stellt die konzeptionelle Vorgehensweise bei der Verwendung von sich zwischen Umweltkompartimenten inhaltlich überschneidenden Listen am Beispiel der Landnutzungsart dar.

Es wird in solchen Fällen eine Hauptliste mit der Obermenge der Listenelemente (hier „Landnutzungsart“) erstellt und die jeweils kompartimentspezifischen Listen davon abgeleitet. Dabei definieren die kompartimentspezifischen Listen ihre eigenen Namen (i.S.v. Bezeichnung) für die Listenelemente, um den Anforderungen des jeweiligen Kompartiments Rechnung tragen zu können.

Ähnlich wird im Teildatenmodell der Stammdaten („master_data“) bei der Modellierung der Messstellen vorgegangen. Die Messstelle fungiert hier als Haupttabelle für alle Messstellen und enthält einige Attribute (UBA-Code, Name, Länder-Code, Bundesland, topographisches Kartenblatt TK25, Koordinaten etc.), die bei den Messstellen aller Umweltkompartimente benötigt werden. Die Messstellentabellen der vier Umweltkompartimente sind mit der Haupttabelle verknüpft und ergänzen die im jeweiligen Kompartiment zusätzlich benötigten Attribute.

Bei der Modellierung von Parametern und Messgrößen¹⁴ war es notwendig, nutzergruppen-spezifische Messgrößenamen und Einheiten zu ermöglichen. Dazu wurde eine zentrale Tabelle Parameter mit den wesentlichen Attributen Parametername (i.S.v. Name der Messgröße), physikalische Größe¹⁵ und Parametereinheit vorgesehen. Hier ist jede Messgröße mit einer FIS Wasser-Boden-Einheit je physikalischer Größe abgelegt. Bei der Ablage von Messwerten wird auf diese Parameter referenziert.

Von den Parametern werden nutzergruppenspezifische Parameteralias abgeleitet, bei denen sowohl der Name als auch Einheit und bei den Wasserkompartimenten die LAWA-Nummer individuell festgelegt werden können. Für den Import von Messwerten werden die in den Parameteraliasen festgelegten Einheiten akzeptiert und die nutzergruppenspezifischen Aliasnamen für die Messgrößenbezeichnungen verwendet. Dabei wird bei abweichenden Einheiten der Messwert passend zum korrespondierenden Parameter umgerechnet.

Somit kann jedes Fachgebiet für seine Datenlieferant:innen die etablierten Namen und Einheiten behalten und trotzdem eine kompartimentübergreifend einheitliche Ablage erreicht werden. Gleichzeitig ist dies eine wichtige Voraussetzung für die Aufnahme weiterer Umweltkompartimente (Meeresumweltdaten, MUDAB) mit voraussichtlich eigenen Anforderungen an Parameternamen und Einheiten.

Es existiert folglich für jeden nutzergruppenspezifischen Parameteralias zuerst ein FIS Wasser-Boden-Parameter, der in allen Umweltkompartimenten zur Verfügung steht und von dem ein nutzergruppenspezifischer Parameteralias abgeleitet werden kann¹⁶.

Boden und Seen haben im Datenmodell eine Sonderstellung. Die Seen, da sie über eine eigene Klasse zur Beschreibung des Sees verfügen und Boden, da die Untersuchungsfläche über eine sehr hohe Zahl an Attributen beschrieben wird und mit den Bodenhorizonten zusätzlich über

¹⁴ Eine Messgröße ist z. B. pH-Wert, Temperatur, Stoff X, Y und Z. Unter Parametern verstehen wir die Kombination aus einer Messgröße, einer Einheit und einer Matrix.

¹⁵ Es werden Messwerte, z. B. von Schwermetallen, sowohl als Massenkonzentration (z. B. mg/l) als auch als Relativgröße (z. B. µg/kg bei Biota) übermittelt. Dabei weicht die Dichte von Biota ggf. von 1 kg/l ab, weshalb eine Einheit je Messgröße nicht reicht, da ohne Dichte nicht umgerechnet werden kann.

¹⁶ Es kann jede Nutzergruppe neue Parameter anlegen (sofern die Kombination aus Messgröße und physikalischer Einheit noch nicht vorhanden ist), um individuelle Parameter davon abzuleiten.

eine vertikale Beschreibung des Bodens verfügen, die ebenfalls eine sehr große Zahl von beschreibenden Attributen aufweist.

Da der Bodenteil der Anwendung zur Aufnahme aller bei UBA vorliegenden Boden-Dauerbeobachtungsdaten dient, kann der Attributumfang nicht oder nur bei bislang vollständig ungenutzten Attributen reduziert werden. Folgende grundlegende Änderungen wurden jedoch vorgenommen:

Die bislang in zwei unterschiedlichen, sich ergänzenden Tabellen beschriebene Untersuchungsfläche wird in einer Tabelle beschrieben. Dies wurde auch für die Bodenhorizonte so umgesetzt. Hintergrund ist, dass unbedingt Duplikate¹⁷ vermieden werden müssen, die durch die bisherige Trennung auftreten können. Weiterhin widersprüche die weitere Trennung dem Stand der Technik. In der Oberfläche sind bodenkundlich zusammenhängende Attributsätze („Titeldaten“, nun „Stammdaten“, „Aufnahmesituation“, „Profilkennzeichnung“ und „Horizonte“) mit den bekannten Bezeichnungen belegt.

Die in der Leistungsbeschreibung im konzeptionellen Modell angedeutete Trennung in thematische Datenbanken wurde anders umgesetzt. Es werden nur die fachlich notwendigen, durch unterschiedliche Inhalte (Strukturen) zwangsweise zu trennenden Inhalte aufgeteilt. Dies geschieht innerhalb einer Datenbank – der des WaBo-Intranet.

Es gibt folglich ein Kernmodell mit Parametern, Messstellen, Messungen und Messwerten, die jeweils jenen, den meisten oder allen Umweltkompartimenten gemeinsamen, Attributumfang umfassen und dazu thematische Erweiterungen je Umweltkompartiment mit spezifischen Attributen (i.S.v. Modulen). Dies ist einerseits IT-technisch geboten, aber auch zweckmäßig, wenn Daten per Indexer oder Suchmaschine kompartimentübergreifend bereitgestellt werden sollen.

Die thematische Trennung erfolgt technisch durch eine Verknüpfung der Messstellen mit den Umweltkompartimenten. Darüber sind auch Messungen und Messwerte jeweils mit einem Umweltkompartiment verknüpft.

¹⁷ Es wurden bei der Sichtung der Bodendatenbank zum Zweck der Konzepterstellung solche Duplikate gefunden, die eindeutig auf die Bereitstellung von Bodendaten direkt auf der Datenbank durch Dritte zurückzuführen sind.

12 Nachrichtenmodelle

Nachrichtenmodelle definieren die Struktur ausgetauschter Nachrichten. Es sind folglich die Datenmodelle für ausgetauschte Daten.

FIS Wasser-Boden sieht in seinen Prozessen den Austausch verschiedener Nachrichten vor. Die für das Projektziel des automatisierten Datenaustauschs wichtigsten Nachrichtenmodelle betreffen die Lieferung von Messdaten der Bundesländer (oder von Organisationen in deren Auftrag) an das UBA. Das für den standardisierten Datenaustausch in der Bundesrepublik verbindliche XÖV sieht für die Übertragung von Daten generell das XML-Format und für die Definition der Strukturen von zu übertragenden Daten die Verwendung von XML-Schemata vor. Daher wurde für die Modellierung und die Visualisierung XML-Schema verwendet.

Für die Differenzierung von Messwerten innerhalb vs. außerhalb der Bestimmungsgrenzen wurde der Choice-Kompositor (Auswahl genau eines Elements) zur eleganten Modellierung verwendet. Generell ist das Design möglichst „flach“, im Sinne von nicht so stark geschachtelt, gehalten. Dies wird insbesondere im Aufbau der XML-Dateien mit den Messwerten sichtbar.

Bezeichnungen wurden, wo fachlich nicht sinnentstellend, einheitlich gewählt, z. B. „Matrix“ als Ersatz für Probenaufbereitung und Trennkennzahl.

Derzeit sehen wir die Matrices Schwebstoff/Sediment, Gesamtgehalt, Probe 2h abgesetzt, gelöste Phase, Fisch und Muscheln vor. Messungen von z. B. Schwermetallen in Fischfilet oder Muscheln werden dementsprechend unter Angabe der Matrix Fisch bzw. Muscheln übermittelt.

In Abstimmung mit dem Auftraggeber wurden Nachrichtenmodelle für den Import von Messwerten entwickelt. Die selten notwendige Aufnahme neuer Messstellen oder die Korrektur an deren Daten ist über entsprechende Formulare in der Benutzungsoberfläche von FIS Wasser-Boden realisiert. Im Fall von (selten erwarteten) großen Zahlen von Messstellen werden diese von einem Dienstleister oder Verfahrenshersteller direkt in die Datenbank übertragen.

Die realisierten Importmodule für Wasser- und Bodendaten orientieren sich an den derzeitigen, realen Lieferbedingungen und verarbeiten die Tabellenformate .xlsx (Excel) und CSV (comma separated values), die direkt aus den in XML-Schema modellierten Nachrichtenmodellen abgeleitet wurden.

12.1 Fließgewässer

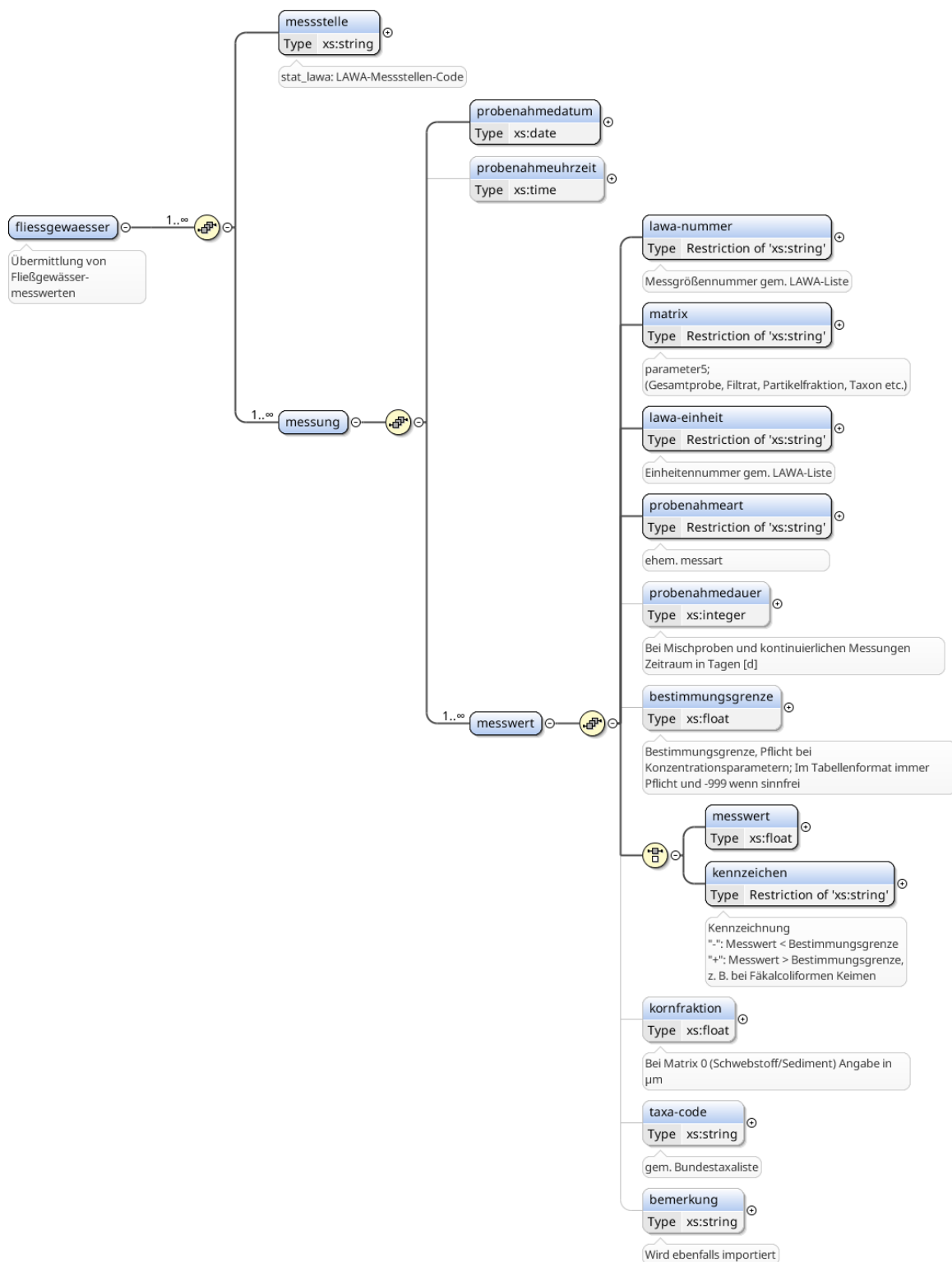
Mit den im Projekt bereit gestellten Informationen wurde ein Nachrichtenmodell für den Import von Fließgewässermesswerten erstellt und als XML-Schema modelliert. Es stellt den Ausgangspunkt für die Entwicklung der Nachrichtenmodelle für die anderen Umweltkompartimente dar, die einen sehr ähnlichen Aufbau aufweisen.

Bei der Entwicklung wurden die in den XÖV-Standards XUBetrieb und XUKommunalabwasser definierten Bezeichner und Strukturen verwendet, die in diesem Fall nutzbar waren¹⁸.

Auf diejenigen Elemente der XÖV-Standards, die die Implementation auf der Seite der Datenlieferanten unnötig erschweren, wurde im Sinne der Zielerreichung verzichtet.

¹⁸ XUBetrieb und XUKommunalabwasser sind XÖV-Standards für den Austausch betrieblicher Umweltdaten und zielen auf die Übertragung von Daten der 11. BImSchV, PRTR und Kommunalabwasserdaten gem. 91/271/EWG ab. Sie weisen nur sehr geringe Überschneidungen mit den in FIS Wasser-Boden zu übertragenden Nachrichten auf.

Abbildung 21: Nachrichtenmodell Messdatenübermittlung Fließgewässer



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Im Rahmen der auf die Konzeptphase folgenden Entwicklung wurden für die Elemente **einheit** und sowie alle die Messgröße definierenden Elemente (**lawa-nummer**, **messgroessencode** etc.) des Schemas Prüfungen für die Einhaltung erlaubter Werte entwickelt. Diese Prüfungen werden auch für die Prüfung übergangsweise weiter verwendeter Tabellenformate benötigt. Dadurch

werden die Schemata übersichtlich gehalten und es wird ermöglicht, Datenlieferanten spezifischere Meldungen zu geben, als dies mit einer XML-Schema-Validation möglich ist.

Mit der separaten Tabellendatei `importtabelle_wasser_konzept.xlsx` sollen Messwerte für Fließgewässer, Seen und Grundwasser importiert werden. Sie enthält ein Tabellenblatt mit Hinweisen und ein Tabellenblatt mit einer Vorlage für die Werte.

12.2 Biologische Parameter

Sowohl für die Fließgewässer als auch die im Text folgenden Seen konnte geklärt werden, ob eine Übertragung und Ablage biologischer Parameter und der zugehörigen Werte möglich ist. Dies kann durch die Einführung einer Taxon-Nummer zur Definition der — hier biologischen — Messgröße erreicht werden. Nachfolgend werden alle derzeit bekannten Fälle ggf. aufzunehmen der biologischer Parameter bzgl. der Art der Aufnahme in das FIS Wasser-Boden diskutiert.

12.2.1 Taxonomie und Abundanz

Die Taxon-Nummer aus der Bundestaxaliste der Gewässerorganismen Deutschlands ist bzgl. des Taxons UND der taxonomischen Kategorie (Stamm, Klasse, Ordnung, Familie, Gattung, Art) eindeutig. Durch die Einführung der Taxon-Nummer (`taxon-nr`) können die sich mit den LAWA-Stoffnummern überschneidenden Taxon-Nummern beim Import von ersteren unterschieden werden.

Ein Taxon wird bzgl. des Namens und der Nummer wie die sonstigen Parameter behandelt. (Die Taxa-DV-Nummern haben den angenehmen Vorteil, eindeutig zu sein i.d.S., dass sich eine Nummer einer Art nicht wieder in Gattung oder Familie findet und umgekehrt.)

Da die Individuen also gezählt werden, ist die Einheit dann Anzahl und zunächst dimensionslos, der Wert muss eine nicht-negative Ganzzahl sein.

Die Matrix soll im Fall Taxon/Abundanz leer bleiben.

Ob die Individuen pro Volumen, pro Fläche oder auf einer Länge gezählt werden, kann der übermittelten Einheit entnommen werden:

- ▶ 1/Volumen (z. B. $1/\text{cm}^3$)
- ▶ 1/Fläche (z. B. $1/\text{m}^2$) oder
- ▶ 1/Länge (z. B. $1/\text{m}$).

12.2.2 Makrophytenausdehnung (max. Tiefe)

Hier führen wir einen Parameter Makrophytenausdehnung mit der Einheit Meter ein. Mit diesem wird die von der Wasseroberfläche Richtung Erdmittelpunkt gerichtete Tiefe in Metern übermittelt (bis zu der Makrophyten gefunden werden).

Die Matrix bleibt wiederum leer.

12.2.3 Chla- und Phaeophytindaten (Konzentrationen) (Zellpigment)

Wir führen zwei neue Parameter mit diesen Namen ein und verlangen als Wert eine nicht-negative Zahl. Die physikalische Größe ist Massekonzentration, die Einheit folglich mg/l o.ä..

Hier sollte voraussichtlich eine Matrix angegeben werden.

12.2.4 EQRs und Messungen (bspw. Schwermetalle, org. Schadstoffe) für die Kompartimente Sediment, Schwebstoff und diverse Organismen

EQRs sind „Ecological Quality Ratios“ und dimensionslose Kennzahlen mit einem Wertebereich von 0 bis 1. Folglich ist die Einheit dimensionslos und der Wert eine nicht-negative Fließkommazahl von 0 bis max. 1.

Es wird zu jeder EQR einen Namen geben, der als entsprechender Parameter aufgenommen wird. Die physikalische Größe ist: dimensionslose Größe, Einheit: dimensionslos.

12.2.5 Zustandsklassen von 1 bis 5 und chemischer Zustand gut / nicht gut

Wenn es mehrere, voneinander zu differenzierende Zustandsklassen gibt, ist für jede Zustandsklasse ein Name zu bestimmen und als Parameter aufzunehmen. Zu diesem ist der Zustandswert zu übermitteln.

Die physikalische Größe ist: dimensionslose Größe, Einheit: dimensionslos.

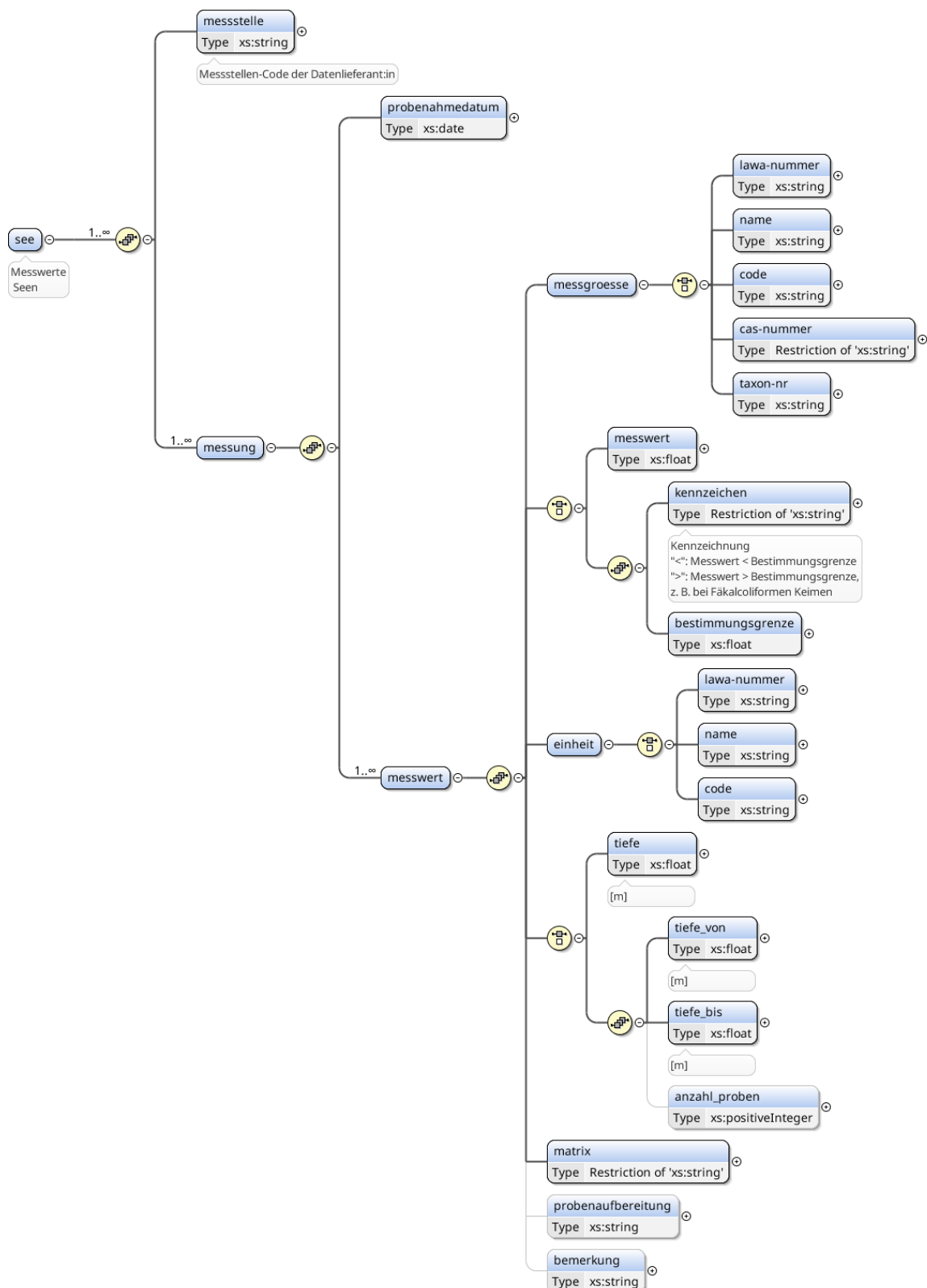
Die Matrix bleibt leer.

Analog kann mit dem chemischen Zustand verfahren werden. Es steht folglich der zukünftigen Aufnahme biologischer Parameter in das FIS Wasser-Boden nichts entgegen.

12.3 Seen

Das hier für die Übermittlung von Seenmesswerten entwickelte Nachrichtenmodell ist dem für Fließgewässer und Grundwasser im Aufbau sehr ähnlich. Hier kann jedoch auf einige fließgewässerspezifische Elemente (XML-Begriff für Attribute) verzichtet werden, hingegen werden zusätzlich Tiefenangaben benötigt. Auf die Aufnahme der interpretationsabhängigen Messzonen („oberflächennah“, „Epilimnion“, „Metalimnion“, „Hypolimnion“ etc.) wurde verzichtet, da sie für eine wissenschaftliche Auswertung nicht herangezogen werden können.

Abbildung 22: Nachrichtenmodell Messdatenübermittlung Seen



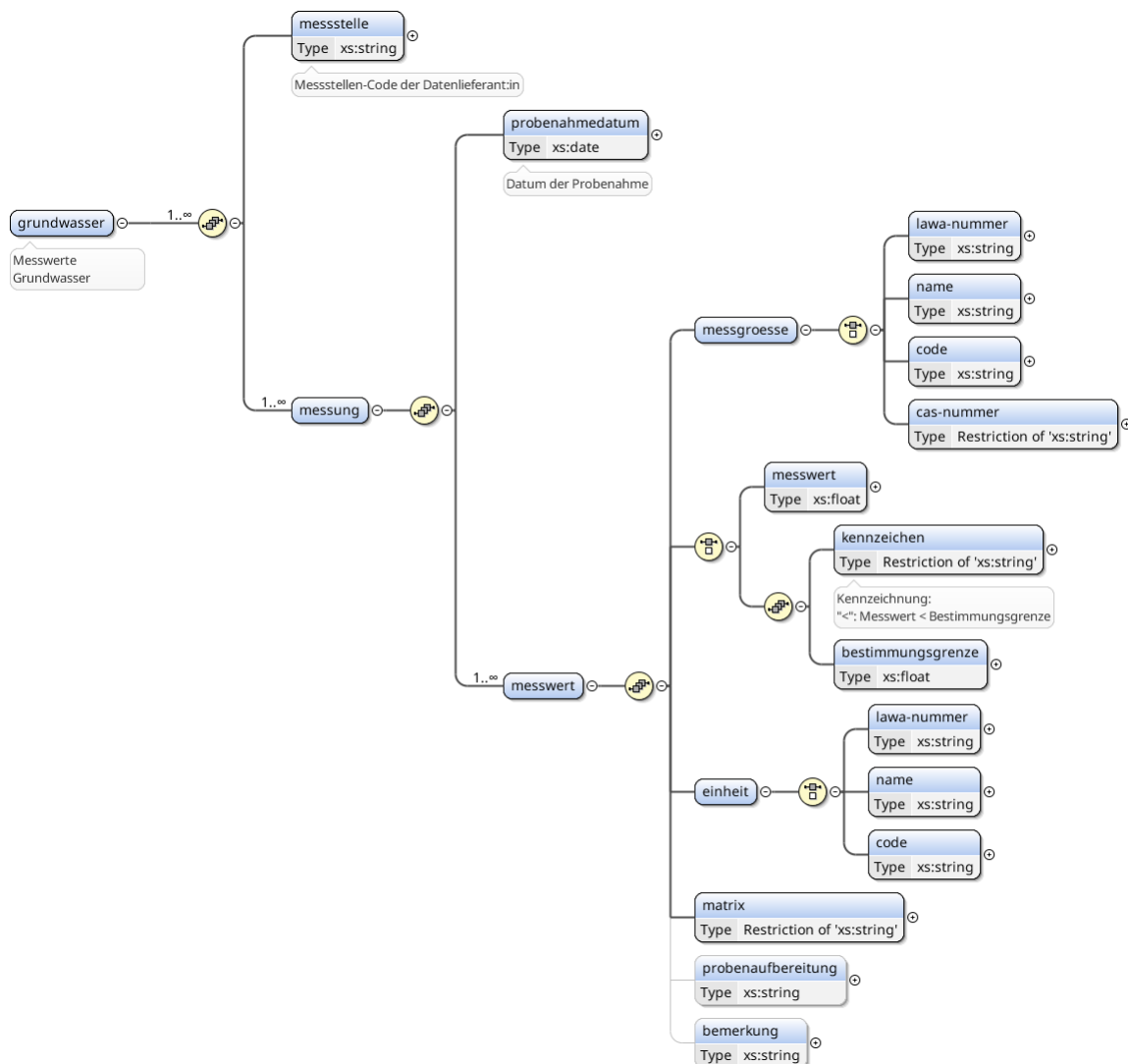
Quelle: eigene Darstellung, ENDA GmbH & Co. KG

12.4 Grundwasser

Für die Übermittlung von Grundwassermesswerten kann das Schema auf die Gemeinsamkeiten der Fließgewässer- und Seen-Schemata (ohne Taxa) reduziert werden. Wie bei den Seen, ist die

Angabe der Bestimmungsgrenze nur dann Pflicht, wenn der Messwert außerhalb der Bestimmungsgrenze des jeweiligen Labors liegt.

Abbildung 23: Nachrichtenmodell Messdatenübermittlung Grundwasser

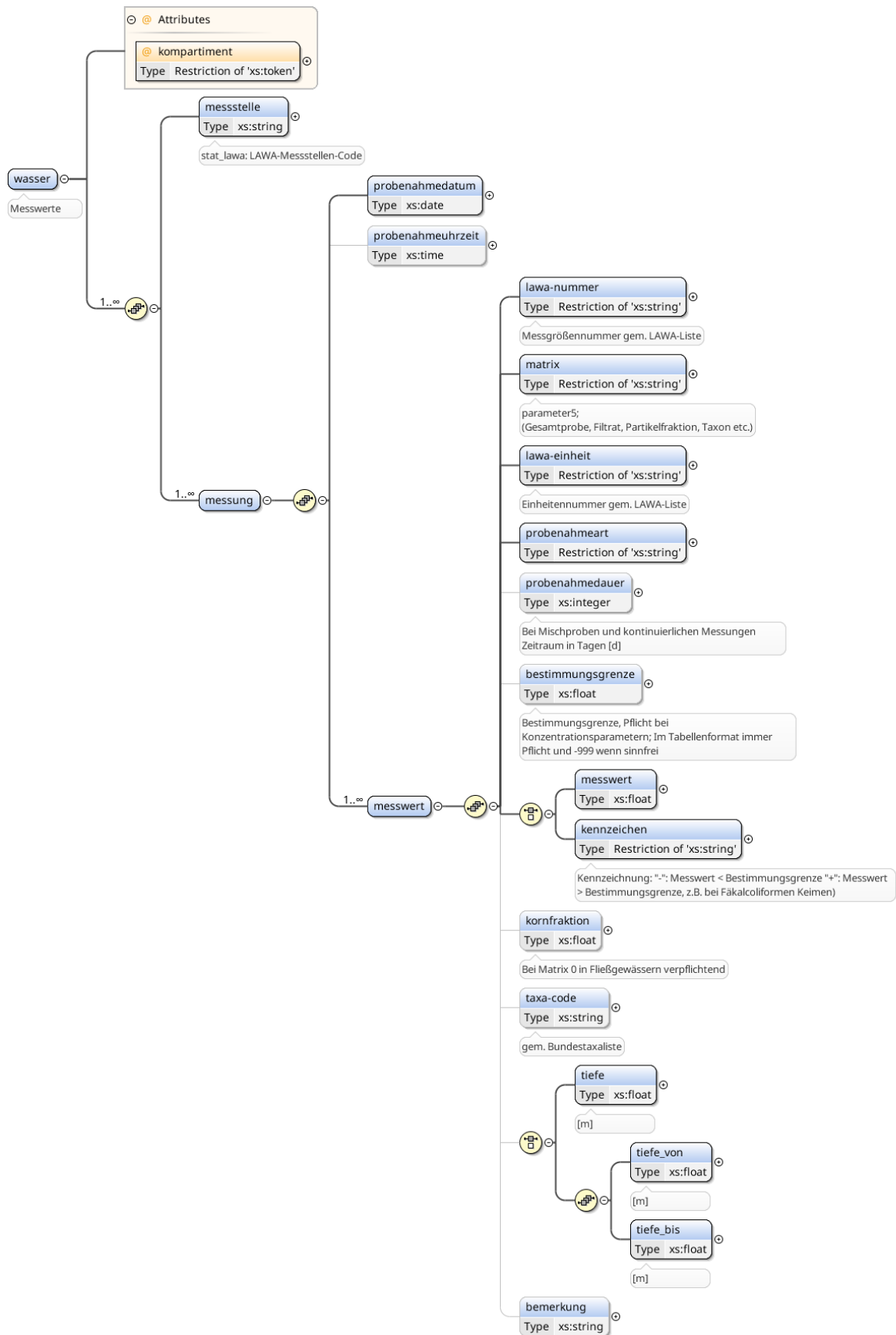


Quelle: eigene Darstellung, ENDA GmbH & Co. KG

12.5 Gemeinsames Nachrichtenmodell Wasser

Die Entwicklung und Abstimmung der spezifischen Import-Nachrichtenmodelle der Wasser-Umweltkompartimente erfolgte iterativ. Die Änderungen und Erweiterungen wurden nach jedem Schritt in ein gemeinsames Nachrichtenmodell Wasser integriert und dieses an die verantwortlichen Fachbetreuenden zurückgespielt. Dadurch konnte erreicht werden, dass alle Wasser-Umweltkompartimente dieses Nachrichtenmodell importseitig verbindlich einsetzen.

Abbildung 24: Gemeinsames Nachrichtenmodell Wasser

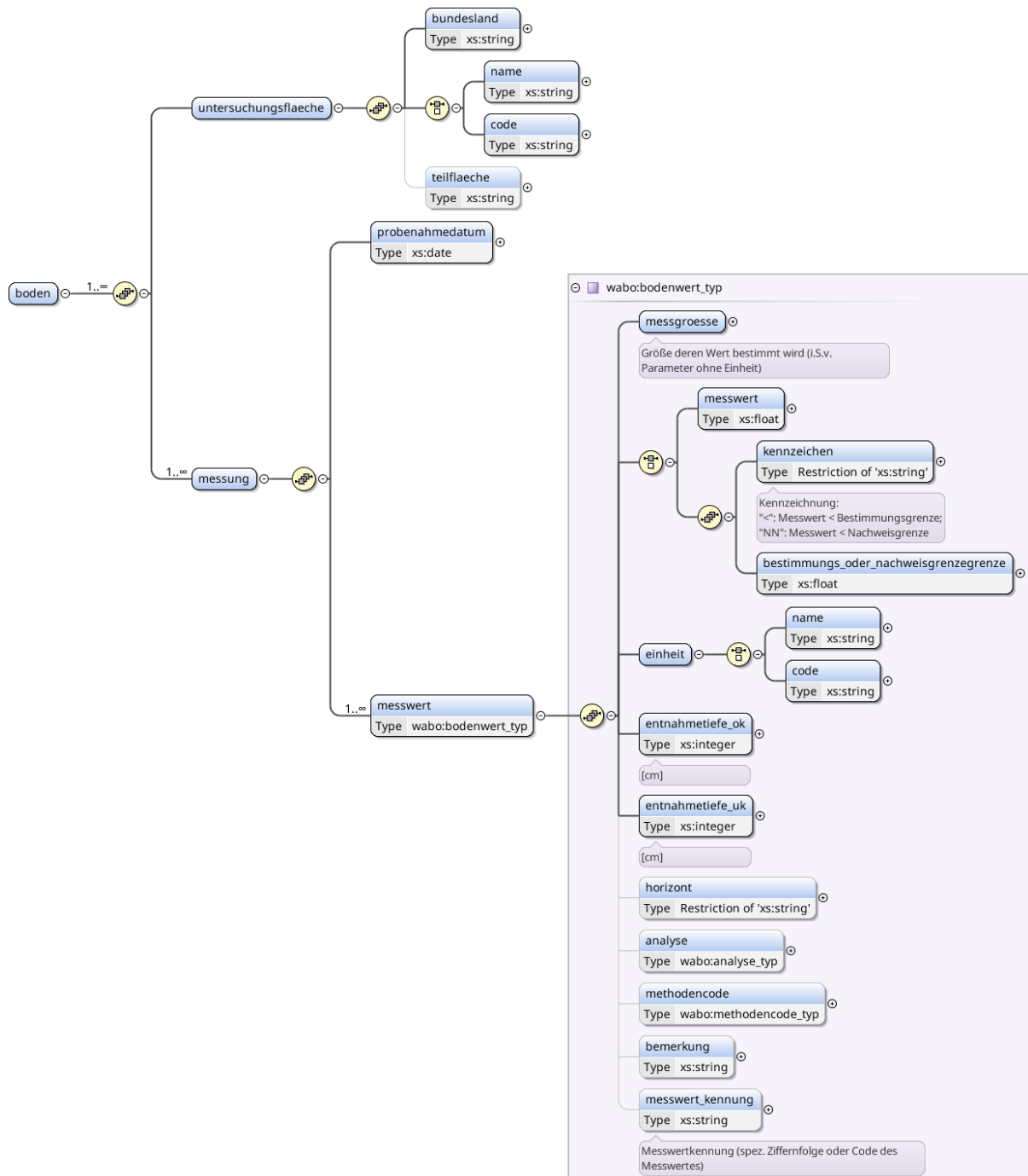


Quelle: eigene Darstellung, ENDA GmbH & Co. KG

12.6 Boden

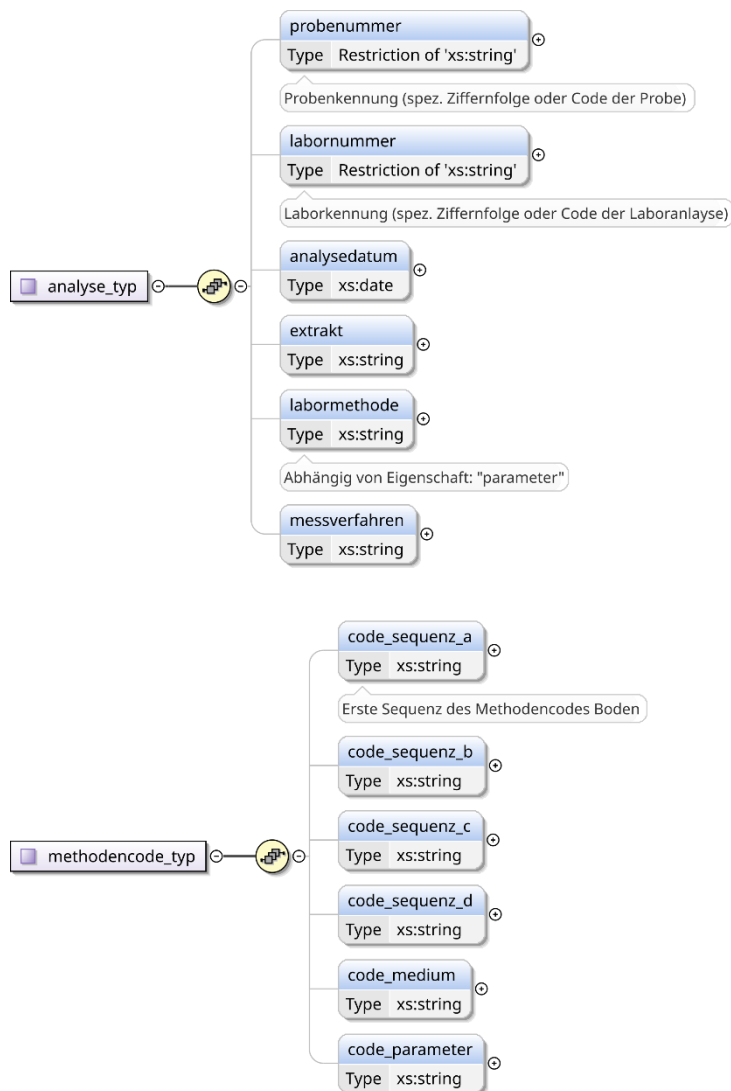
Das Schema für die Übermittlung von Bodenmesswerten ist mit größtmöglicher Ähnlichkeit zu den Gewässer-Schemata entwickelt worden und greift weitgehend den Modellierungsumfang aus dem neuesten, kompakten Format für die Beschreibung von Messdaten zum Zwecke des Datenaustauschs aus dem Bodenschutz auf.

Abbildung 25: Nachrichtenmodell Messdatenübermittlung Boden, Hauptmodell



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

Abbildung 26: Nachrichtenmodell Messdatenübermittlung Boden, Analyse und Methodencode



Quelle: eigene Darstellung, ENDA GmbH & Co. KG

12.7 Tabellenformate

Es werden derzeit noch Tabellenformate benötigt, die die Lieferung der oben definierten Nachrichtenformate in Form von XML-Schemata ersetzen, um den realen Anforderungen und Möglichkeiten der Datenlieferant:innen zu entsprechen. Nach fachlicher Kontrolle und Freigabe der Schemata sind diese Tabellenformate in enger Anlehnung an die Schemata (Elementnamen werden Spaltenköpfe etc.) in enger Abstimmung mit dem UBA entwickelt worden. Sie sind im Dokumentations- und Informationsmodul für die Datenlieferant:innen zum Download bereitgestellt. Die Vorlagedateien enthalten jeweils ein Tabellenblatt mit Metainformationen und detaillierten Erläuterungen, ein Tabellenblatt als Vorlage für die zu übermittelnden Daten und ein Tabellenblatt mit Versionsinformationen.

13 Data Warehouse

Das für das FIS Wasser-Boden erstellte Data Warehouse ist ein Datei-Speicher. Ein Eintrag im Datawarehouse besteht aus

- ▶ einer Datei mit Dateinamen
- ▶ einem Titel
- ▶ einer Beschreibung und
- ▶ Metadaten (Datum; Person, die veröffentlicht hat)

Das Data Warehouse besitzt eine hierarchische Ordnerstruktur. Nutzer des Data Warehouse müssen keinen Zugang mit Passwort besitzen, das es für alle Beschäftigten des UBA mit Zugang zum UBA-internen Netzwerk (Intranet) offen ist. Das Data Warehouse ist von den Fachdaten der Umweltkompartimente getrennt – man erhält darüber keinen Zugang auf die Fachdaten. Weiterhin ist es nur eingeloggten Nutzern möglich, dem Data Warehouse Inhalte hinzuzufügen.

13.1 Befüllung des Data Warehouse und Zugriff auf die Inhalte

Über die Weboberfläche der Anwendung werden alle Ordner der obersten Ebene angezeigt. Diese umfassen mindestens die vom System selbst benötigten Ordner „Berichte“ und „Abfragen“.

- ▶ **Berichte:** Hier werden die durch die verschiedenen Berichte erzeugten Berichtsdateien hinterlegt. Wird ein neuer Bericht erzeugt, mit Namen versehen und konfiguriert, wird im Data Warehouse unterhalb vom Ordner „Berichte“ ein Ordner mit dem Namen des Berichts angelegt. In diesem werden alle mit dem Bericht erzeugten Berichtsdateien, in der Regel Tabledateien im .xlsx-Format, hinterlegt. Ihr Titel enthält den Berichtsnamen und einen Zeitstempel.
- ▶ **Abfragen:** Mit dem Abfragewerkzeug konfigurierte Abfragen können zyklisch automatisiert ausgeführt werden. Die Ergebnisse dieser automatisierten Abfragen werden in einem Ordner mit dem Namen der Abfrage innerhalb des Ordners Abfragen abgelegt. Auf diese Weise können ausgewählte und qualitätsgesicherte Fachdaten in regelmäßigen Intervallen zur weiteren Verarbeitung anderen UBA-Mitarbeitenden zur Verfügung gestellt werden. Das bei der Konfiguration der jeweiligen Abfrage festgelegte Ausgabeformat (.xlsx, CSV oder JSON) wird auch für die Datei im Data Warehouse verwendet.
- ▶ **Per manuellem Upload:** So können von angemeldeten Nutzenden beliebige Dateien, also auch Text-Dokumente (PDF, .docx etc.) oder Grafiken, hinterlegt werden.

Der Einsatz einer Volltextsuche wurde im Kapitel 15, „Integrierte Forschungsaspekte und Umsetzung der Forschungsfragen“, untersucht.

Eine Web-API zur Nutzung der Suchmaschine halten wir für nicht sinnvoll. Eine technische Realisierung wäre jedoch möglich. Hintergrund ist, dass APIs in der Regel für den automatisierten Austausch von Daten vorher bekannter Struktur und meist auch vorher recht genau bekanntem Umfang verwendet werden. Ein Suchmaschinenergebnis wird (oder sollte zumindest) nicht Teil eines solchen Datenaustauschprozesses sein, da sein Ergebnis bei Änderung des Suchbegriffs weder in Umfang noch in Struktur vorab bekannt ist.

Bezüglich des Designs ist sowohl ein Zugang aus der Fachanwendung FIS Wasser-Boden ohne erläuternde oder das Projekt beschreibende Inhalte implementiert, als auch ein separater

Zugang zu den Inhalten des Data Warehouse ohne Authentifizierung. Der Zugang zu den Fachdaten des FIS Wasser-Boden soll nur über die Inhalte des Data Warehouse erfolgen, um den Fachbetreuenden die volle Kontrolle über die zugänglichen Daten zu geben. Die Nutzung des Abfragewerkzeugs kann bei Zugriff auf nicht qualitätsgesicherte Daten durch UBA-Mitarbeitende außerhalb der verantwortlichen Fachgebiete zu Irritationen führen. Dies soll vermieden werden.

14 Prüfung der Einbindung weiterer Umweltmedien

14.1 Meeresumweltdatenbank (MUDAB)

Die Daten der MUDAB und die mit der MUDAB bereit gestellten Produkte sollen zukünftig in das FIS Wasser-Boden überführt und durch dieses zur Verfügung gestellt werden. Hier wurde geprüft, wie diese Integration durchgeführt werden kann, welche Anpassungen durchzuführen sind und mit welchen Aufwänden dabei zu rechnen ist.

Die Integration der Meeresumweltdatenbank (MUDAB) wurde bei der Erstellung des Konzepts so berücksichtigt, dass diese Erweiterung bei der Integration von Messgrößen und Einheiten (Parametern) keine Probleme verursachen wird. Die derzeit von MUDAB bereit gestellten Projektstationen weisen allerdings einen geringeren Informationsgehalt auf, als die derzeitige Haupttabelle der FIS-WaBo Messstellen¹⁹. Hier wird entweder ein Bundesland für die MUDAB-Messstellen gefunden oder die Angabe des Bundeslandes der Messstelle wird optional gemacht.

Ursprünglich lag die MUDAB Datenbank nicht vor, so dass der Datenumfang aus der Webseite der MUDAB geschätzt und mit zunächst nur sichtbaren 200.000 Stationsparametern deutlich unterschätzt wurde.

Tatsächlich weist die aus Oracle exportierte und vom Auftragnehmer in Oracle reimportierte MUDAB in zwei Tabellen jeweils etwas mehr als fünf Millionen Datensätze auf. Das sind derzeit ungefähr so viele, wie die derzeitigen Kompartimente des FIS Wasser-Boden zusammen enthalten. PLC-Daten²⁰ konnten in der Oracle-MUDAB nicht explizit identifiziert werden.

14.1.1 Zusammenfassung

Die MUDAB Oracle-Datenbank entspricht im Stil nicht der seit den 2000er Jahren üblichen Modellierung von Daten und entscheidende Teile entsprechen auch nicht der Struktur, die für eine wissenschaftliche Nutzung der Daten notwendig ist²¹. Andere Teile, insbesondere die Dokumentation der Attribute der Tabellen, ermöglichen die Interpretation der fachlichen Intentionen. Es werden jedoch bei dem derzeitigen Stand verfügbarer Dokumentation und Unterlagen erhebliche Aufwände notwendig sein, die Inhalte zu analysieren und so aufzubereiten, dass sie in eine nutzbare Form gebracht und in das FIS Wasser-Boden übertragen werden können²².

Nach den vorliegenden Daten und auch den Erkenntnissen aus dem Gespräch mit BfG ist davon auszugehen, dass FIS Wasser-Boden sowohl bzgl. des Gebietes der Probenahme als auch bzgl. des Zeitraums der Probenahme erweitert werden muss. Der Zeitraum der Probenahme kann sich voraussichtlich auch auf Messkampagnen²³ beziehen, die dann ebenfalls in FIS Wasser-Boden Eingang halten würden.

¹⁹ Die MUDAB-Messstellen enthalten die Koordinaten der Probenahme

²⁰ PLC — Pollutant Load Compilation; Eventuell lassen sich die für PLC benötigten Daten aus anderen Informationen der MUDAB ableiten.

²¹ Es fehlen zahlreiche Regeln zur referenziellen Integrität für die Listen bzw. Kataloge. Die Semantik der Tabellen- und Spaltenbezeichner ist i.d.R. nicht ersichtlich, da Abkürzungen benutzt wurden oder Spalten mit dem Namensteil "ID" versehen wurden, die keine Referenzen enthalten, sondern Namen. Beschreibungen zu weiteren Interpretationsproblemen finden sich im folgenden Text, Abschnitt Messwerte.

²² Für eine Integration der MUDAB-Inhalte in das FIS Wasser-Boden kann die MUDAB nicht geändert werden, sondern die für die Aufnahme der Daten noch fehlenden Tabellen und Attribute werden im FIS-WaBo ergänzt und nachfolgend die MUDAB-Inhalte per Programm übertragen. Ansonsten wären die bereits in FIS-WaBo implementierten Werkzeuge nicht für die MUDAB nutzbar (Abfragewerkzeug, Reporterstellung etc.) und die erhofften Synergieeffekte könnten nicht eintreten.

²³ Unter Kampagnen werden in den Umweltbereichen Grundwasser, Seen und Fließgewässer sowie Boden durch Anfangs- und Enddatum begrenzte Zeiträume, meist Jahre, verstanden, für die bereits erfasste Messwerte bzgl. bestimmter Messgrößen (z. B.

Eine Integration der MUDAB in FIS Wasser-Boden halten wir ungeachtet des Zustands der MUDAB nicht nur für möglich, sondern auch für geboten, damit bei einem Bearbeiter:innenwechsel deren Inhalte weiter interpretiert und genutzt werden können. Dies wird mit der Anwendung FIS Wasser-Boden, die der Architekturrichtlinie IT-Bund und dem XÖV folgt, einfacher möglich sein.

Sofern sich die Berichtsaufwände der MUDAB nach einer Integration in das FIS Wasser-Boden in der Größenordnung der gemeinsamen Aufwände der Berichterstattungen der Wasserkompartimente Fließgewässer, Grundwasser und Seen bewegen, ergibt sich daraus ein erhebliches Einsparpotenzial.

14.1.2 Arbeitspakete zur Integration der MUDAB in FIS Wasser-Boden

Eine erfolgreiche Integration umfasst folgende Arbeitspakete:

- ▶ Modell- und Datenübertragung
 - Analyse der bestehenden Datenbank
 - Übertragung aller Tabellen- und Attribut-Bezeichner in eine Form, konform zu den Anforderungen des XÖV und der Architekturrichtlinie IT Bund
 - Erweiterung des Modellkerns von FIS Wasser-Boden²⁴
 - Entwicklung eines Programms zur Datenübertragung
- ▶ Importschnittstellen und Exportformate (Produkte)
 - Beschaffung von repräsentativen Eingangsdaten²⁵ für alle zukünftig genutzten Input-Kanäle
 - Entwicklung von Importschnittstellen, die auf die jeweilige Aufgabe abgestimmt sind. Hier sind die Aufwände abhängig von Art und Anzahl. Eine grobe Kalkulation würde ca. 8 bis 10 Zeiteinheiten pro Importschnittstelle vorsehen. Strukturell einfache Schnittstellen (beispielsweise nur Messwerte) auf Basis von REST-Interfaces sind eher preiswerter (4 bis 6 Zeiteinheiten) als strukturell komplexe (z. B. expeditionsbeschreibende Strukturen) unter Einsatz eher unzuverlässiger (Excel-)Formate. Referenzszenario: Zwei Importformate (jeweils Messwerte und Expeditionsdaten), einmal Excel, einmal JSON oder XML via REST-Interface mit grundlegender, einmal zu entwickelnder QS.
 - Aufarbeitung aller MUDAB-Produkte aus dem Bereich Datenlieferungen²⁶, die weiterhin bereit gestellt werden sollen i.d.S., dass die Zielformate, -strukturen und -datenumfänge

Schadstoffkonzentrationen, Temperatur, pH-Wert etc.) an bestimmten Messstellen (i.S.v. festen Orten, an denen die Beprobung stattfindet) als Datenlieferung zu übermitteln sind.

Die MUDAB-Ausfahrten beziehen sich nicht auf Datenlieferungen, sondern auf Beprobungszeiträume. Letztere würden zukünftig als Messkampagne bezeichnet und eine entsprechende Struktur im FIS Wasser-Boden würde angelegt, um die Inhalte der MUDAB-Ausfahrten übernehmen und für die Anwender:innen bereitstellen zu können.

²⁴ Dies schließt die ICES Codes (Listen bzw. Kataloge) mit ein, für die ebenfalls semantisch interpretierbare Bezeichner und Erläuterungen generiert werden müssen.

²⁵ Unter repräsentativen Eingangsdaten sind hier Beispiele für diejenigen Dateien gemeint, die derzeit in die MUDAB importiert werden. Repräsentativ meint hier, dass diese Dateien ein großes, möglichst vollständiges Spektrum an Angaben enthalten, so dass es einem Auftragnehmer ermöglicht wird, einen möglichst vollständigen und robusten (i.S.v. datentoleranten) Importer zu entwickeln.

²⁶ Dies sind mindestens die ICES Berichtsformate für Wasser, Sediment und Biota für Nähr- und Schadstoffe, für biologische Effekte, für Benthos und Plankton. Spezifische, sich ändernde Auswertungen sehen wir nicht als Berichtsformat, ebenso GIS-Karten. Inhalte für individuelle Berichte (z. B. Nitratbericht) sollten über gespeicherte und somit wiederverwendbare und anpassbare Abfragen mit

soweit spezifiziert sind, dass Reportwerkzeuge dafür entwickelt werden können. Dazu sind die Entwicklung der entsprechenden Exportformate und ggf. Ansteuerung von automatisierten Schnittstellen vorzusehen. Hier sollte mit drei Zielformaten gerechnet werden.

- ▶ Optional, falls nicht bereits über DataCube o.ä. abgedeckt:
Erstellung eines Konzeptes für eine heutigen Nutzungsstandards entsprechende, informative Webpräsenz, natürlich in enger Abstimmung mit dem UBA und anschließend deren Realisierung.
- ▶ Optional: Erstellung einer Projektwebseite (z. B. Wiki) mit Hilfeinhalten und Erstellung eines oder zweier Videotutorials mit Hinweisen zur Nutzung und Hinweisen auf Fallstricke.
- ▶ Projektkoordination und -dokumentation

14.1.3 Detaillierung der Arbeitsschritte

Die Erweiterung des Modellkerns von FIS Wasser-Boden sollte so erfolgen, dass neue Inhalte unter Einbeziehung aller nutzbaren, bestehenden Teile von FIS Wasser-Boden integriert werden, so dass das Ergebnis wieder homogen ist. Damit sind die bestehenden Werkzeuge (z. B. Abfragewerkzeug, Data Warehouse) ohne großen Aufwand auch für die neu integrierten Daten nutzbar, FIS Wasser-Boden bleibt wie gefordert erweiterbar und die Pflegeaufwände eskalieren nicht.

14.1.3.1 Importschnittstellen

Zur Thematik des bisherigen Datenimports in die bestehende MUDAB sind folgende Informationen übermittelt worden:

- ▶ Es findet ein Datenimport statt
- ▶ Es werden Excel-Tabellen auf das Web-Portal WasserBLiCK hochgeladen.
- ▶ Zur Prüfung der Daten wird ein mit hohen Lizenzkosten behaftetes Programm „Feature Manipulation Engine“ (ein ETL-Werkzeug) eingesetzt
- ▶ Es werden keine Crawler eingesetzt, die von verschiedenen Quellen (wie FTP-Laufwerken) Daten einsammeln

Das deutet auf teils individuelle Lösungen mit häufig notwendiger manueller Nachbearbeitung hin und vermittelt nicht den Eindruck von Stringenz. Unter solchen Bedingungen sind hohe Pflege- und Supportaufwände zu erwarten.

Hier ist folglich zu entscheiden, welche Produkte zukünftig unterstützt werden sollen und ob ggf. für komplexe Datenstrukturen Web-Frontends für die Datenlieferanten bereitgestellt werden sollen. Hintergrund ist, dass die Bereitstellung komplexerer Daten mit Excel recht fehleranfällig bzw. eher aussichtslos ist und eventuell Datenlieferanten nicht in der Lage sein werden, XML- oder JSON-Dateien komplexerer Daten automatisiert (und konform zum vereinbarten Nachrichtenmodell) zu erstellen, um diese an eine Importschnittstelle zu übergeben.

Vom Einsatz proprietärer, mit hohen Lizenzkosten behafteter Werkzeuge raten wir ab. Eher sollte eine wiederverwendbare, i.S.v. mit wenig Aufwand auch für andere Projekte des UBA nutzbare Lösung geschaffen werden, die keine Lizenzkosten verursacht und die durch UBA in dessen Sinne beliebig gesteuert werden kann.

dem Abfragewerkzeug dem System entnommen werden. Die Ergebnisse sind in den tabellarischen Formaten .xlsx und CSV herunterladbar.

Eine dem ggf. widersprechende Anforderung des UBA ist jedoch, die Upload-Workflows nicht zu verändern, da die bestehenden Workflows auf Basis von Excel fest etabliert sind und ihrerseits auf IT-Prozesse bei den Datenoriginatoren aufsetzen und Mehraufwand für die Datenoriginatoren vermieden werden soll. Parallel implementierte, moderne Workflows (z. B. auf Basis von REST-Schnittstellen) würden jedoch vom UBA begrüßt.

Die bzgl. der Kosten und der Zuverlässigkeit beste Kombination von Methoden für die Aufnahme von Daten ist die Bereitstellung von Formularen für sich selten oder nie ändernde und komplexe Daten (Beschreibung von Messstellen, Gebieten, Expeditionen etc. mit ggf. mehreren, in Relation stehenden Objektklassen) in Verbindung mit einer automatisierten Importschnittstelle für Massendaten (Messwerte). Letztere ist in der Regel um eine strukturell analoge Schnittstelle für den Import von Tabellendaten (Excel oder CSV) zur Unterstützung derjenigen Datenlieferanten zu ergänzen, die keine Mittel für einen Programmierer und ggf. auch keine geordnete Datenhaltung haben.

Sollen manuelle Bearbeitungen von Datenlieferungen bei UBA oder einem Dienstleister vermieden werden, ist es unerlässlich, das Format für jeden Datenimport klar vorzugeben, eine Zusage der Aufnahme von Daten ausschließlich für in diesem Format gelieferte Daten zu machen und Daten, die das Format nicht einhalten, konsequent zurückzuweisen.

Das UBA kennt derzeit keine gesetzliche Grundlage, die die Anforderung der Daten von den Mee-resanrainer-Ländern und eine klare Spezifikation des Formats ermöglichen würde. Insofern sollte ein Format in Kooperation mit den Ländern entwickelt werden und sowohl Länder als auch UBA eine Vereinbarung zur Nutzung des gemeinsam entwickelten Formats unterzeichnen.

Um die notwendigen Betreuungsaufwände bei UBA bzw. einem beauftragten Dienstleister in Grenzen zu halten, sind Hilfestellungen durch eine Projektwebseite mit Hinweisen, Beispielen und Anleitung sowie Videotutorials sinnvolle Mittel.

14.1.3.2 Aufarbeitung aller MUDAB-Produkte aus dem Bereich Datenlieferungen

Die Aufarbeitung der MUDAB-Produkte ist für alle Datenlieferungen notwendig, die auch zukünftig bereitgestellt werden sollen. Dafür sind die Zielformate, -strukturen und -datenumfänge soweit zu spezifizieren, dass Reportwerkzeuge dafür entwickelt werden können. Anschließend sind diese Reportwerkzeuge zu erstellen, so dass die Datenlieferungen automatisiert durchgeführt werden können.

Wenn das Zielsystem einen manuellen Upload von Daten vorsieht, können die Produkte aus dem FIS Wasser-Boden DataWarehouse heruntergeladen und nachfolgend entsprechend hochgeladen werden.

Existieren empfängerseitig bereits automatisierte Schnittstellen, so sollten diese genutzt werden. In dem Fall würde das FIS Wasser-Boden Reportmodul nach der Freigabe der Daten diese automatisch übermitteln und eine elektronisch ausgestellte Empfangsbestätigung, sofern eine solche vom Empfängersystem bereitgestellt wird, ablegen.

Der ICES sowie Reportnet der EUA führen als Datenempfänger QS-Prüfungen mit Warnungen und Blockern für Datensätze durch. Die Datenabgabe ist folglich ein iterativer Prozess und kann sehr zeitaufwendig sein, mit Telefonaten und Email-Klärungen zu Blockern. Die Datenformate ändern sich bei jeder Berichterstattung etwas. Dieser Prozess ist aus verschiedenen Berichtspflichten (z. B. aus Grundwasser, aber auch der Kommunalabwasserberichterstattung gemäß EU-Richtlinie 91/271/EWG) bereits bekannt. In den entsprechenden Berichtspflichten wurden die manuellen Bearbeitungsaufwände dadurch erheblich reduziert, dass QS-Prüfungen des

Empfängers bereits in die für die Berichterstattung verwendeten Anwendungen übernommen wurden und so ein Großteil der potenziellen Fehler vorab vermieden wurde.

Es liegt auf der Hand, dass die Aufwände stark von der Anzahl der zu unterstützenden Exportstrukturen, also den unterschiedlichen zu erstellenden Reports bzw. Exportschnittstellen und deren jeweiliger Komplexität abhängen.

Auch die Anzahl der einzuhaltenden Regularien und somit der vorher auf den MUDAB-Daten durchzuführenden QS-Prüfungen können stark variieren und damit den Aufwand in die Höhe treiben. Im Rahmen der Interviews zur Prüfung der Einbindung der MUDAB wurde von teilweise mehr als 200 verschiedenen QS-Prüfungen berichtet, die durchzuführen sind, um eine Datenlieferung überhaupt durchführen zu können. Die fehlerfreie Bereitstellung solcher Zielformate ist deutlich aufwändiger, als die Bereitstellung von Daten, die nur wenigen nachfolgenden Prüfungen unterzogen werden.

14.1.3.3 MUDAB Webpräsenz

Die MUDAB Webpräsenz enthält Links auf Berichtsformate. Nach Erstellung eines Zugangs, Rücksprache mit dem Betreiber und dem Einräumen entsprechender Rechte und erneuter Nachfrage mit positiver Antwort, sind diese nicht nutzbar (siehe Anhang F, „MUDAB Bestandsaufnahme“). Dies ist über einen längeren Zeitraum beobachtet und telefonisch kommuniziert worden.

Aus diesem Sachverhalt kann jedoch nicht geschlussfolgert werden, dass die Bereitstellung von Informationen über Berichtsformate überflüssig ist. Es ist auch möglich, dass die Form oder die Inhalte für den Kreis der Nutzenden, die man damit erreichen will, nicht passend oder nicht aktuell genug sind. Beispielsweise ist denkbar, dass die Abstimmung über Berichtsformate auf bilateralen Ebene etabliert ist und funktioniert und dieser Teil der Webpräsenz aus diesen Gründen derzeit nicht benötigt wird.

Sollte von Seiten des zukünftigen Betreibers der MUDAB der Wunsch bestehen, diesen Teil der Webpräsenz zu erneuern, empfehlen wir, vorab den Kreis der potenziell Nutzenden zu identifizieren und mit diesen in Kontakt zu treten. Ziel dieser Kontaktaufnahme sollte sein, die Notwendigkeit zu klären und bei gegebener Notwendigkeit, die Inhalte und die funktionale Gestaltung festzulegen.

14.2 Pflanzenschutzmittel (PSM)

Das Fachgebiet Pflanzenschutzmittel des Umweltbundesamts verfügt über eine Datenbank, die Teile des Aufgabenspektrums im Fachgebiet abdeckt, jedoch nicht in die IT des Umweltbundesamts integriert ist und fachlich nicht an die sich ändernden Bedarfe des Fachgebiets angepasst werden kann. Da sich wesentliche Teile der Aufgaben des Fachgebiets auf die Untersuchung des Grundwassers beziehen und dieses mit FIS Wasser-Boden zur fachlichen Nutzung in einer anderen Abteilung bereitsteht, besteht der Wunsch, zukünftig das FIS Wasser-Boden auf für die Aufgaben des Fachgebiets Pflanzenschutzmittel zu nutzen.

Es sind Gespräche zur Evaluierung der Machbarkeit der Einbindung in FIS Wasser-Boden für den Bereich Pflanzenschutzmittel geführt worden. Auf deren Basis konnte eine grobe Abschätzung der damit verbundenen zeitlich/organisatorischen Aufwände vorgenommen werden, die für eine Entscheidung zwischen einer Integration während der Entwicklungszeit von FIS Wasser-Boden und einer Integration nach Fertigstellung genutzt wurde. Die Entscheidung für eine Integration nach Fertigstellung erfolgte aufgrund noch ausstehender, inhaltlicher Informationen und aufgrund von organisatorischen Randbedingungen.

14.2.1 Tätigkeitsspektren im Bereich PSM

Das Fachgebiet Pflanzenschutzmittel hat folgende, mit der Erhebung und Auswertung von Daten verbundene Tätigkeitsschwerpunkte:

- ▶ Monitoring Of Pesiticides in Ground Water (MOP-GW): Die Überwachung des Grundwassers auf Pestizide
- ▶ Nachzulassungsmonitoring (NZM): Monitoring, vor allem des Grundwassers, nach der Zulassung von Pflanzenschutzmitteln
- ▶ Fundaufklärung (FA): Aufklärung der Ursachen, die zum Fund von Pflanzenschutzmitteln geführt haben
- ▶ Bußgeldbewährte Anwendungsbestimmung NG301-1 zum Schutz von Trinkwasser (NG301)

14.2.2 Herausforderungen durch PSM

Bislang ist der Einsatzzweck von FIS Wasser-Boden als integrierte Vollzugsdatenbankanwendung, die kompartimentspezifisch erweiterbar ist, vorgesehen. Dazu gehören primär die Fähigkeit zur Aufnahme von Daten (Datenimport) sowie die Bereitstellung von Daten auf verschiedenen Wegen (Abfrageergebnisse, Reporterstellung, DataWarehouse).

Für eine wirksame Unterstützung bei der Bearbeitung der Fundaufklärungen von PSM wird jedoch auch eine Prozessunterstützung benötigt. Das FIS Wasser-Boden ist darauf noch nicht ausgerichtet, was jedoch eine IT-technisch lösbare Aufgabe ist. Für die Implementation einer entsprechenden Prozessunterstützung werden die Prozessdokumentationen oder alternativ Prozessaufnahmeprotokolle aus dem Bereich der PSM benötigt. Solche liegen noch nicht vor. Eine textuelle Beschreibung der Arbeitsabläufe, aus der Teile der Prozesse entnommen werden könnten, liegt derzeit ebenfalls nicht vor²⁷. Hier ist folglich eine Prozessanalyse und -dokumentation durchzuführen. Vor dem Hintergrund von gesetzlich geregelten Abläufen ist eine Prozessoptimierung vrs. nur im Sinne der Verbesserung technischer Abläufe durchführbar. Eine Änderung der zwischen dem UBA und dem BMU festgelegten Vorgehensweisen ist voraussichtlich nicht realisierbar.

Die von der Fachabteilung in Form von überlappenden Excel-Tabellenbereichen begonnene Dokumentation der Datenmodelle ist ein Ausgangspunkt zur Erarbeitung eines konsistenten, möglichst aufgabenübergreifenden Datenmodells. Hier sind die Fragen bzgl. der notwendigen Formate zur Aufnahme von Daten zu klären. Ebenso sind die mit oder aus diesen Daten zu generierenden Produkte zu identifizieren und möglichst in die in FIS Wasser-Boden bzgl. der Struktur bereits bekannten Produkte (evtl. Reports, sicher Abfrageergebnisse, voraussichtlich auch Data Warehouse) konzeptionell und später auch IT-technisch zu integrieren. Diese Aufgabe überschneidet sich mit der Prozessanalyse, bei der die Produkte bereits zutage treten sollten.

²⁷ Das alles ist durchaus nicht ungewöhnlich, nur eben für eine belastbare Analyse der Bedarfe ungeeignet.

15 Integrierte Forschungsaspekte und Umsetzung der Forschungsfragen

15.1 Nutzung fremder Webdienste

Hier war zu untersuchen, wie die Einbindung von Webdiensten von Datenbereitstellenden in das Gesamtsystem integriert und wie diese Datensätze zur weiteren Verarbeitung in der gesamten Prozesskette verarbeitet werden können. Dabei sollte berücksichtigt werden, dass es bisher keine einheitliche Standardisierung zur dienstebasierten Datenbereitstellung von Wasser- und Bodenzustandsdaten in Deutschland gibt. Als Ergebnis war ein Vorschlag zu erstellen, wie eine solcher einheitlicher Dienst aufgebaut und technisch strukturiert sein sollte.

Es konnte gezeigt werden, dass ein vom LUNG MV bereit gestellter Web Feature Service (WFS) unter der Internetadresse `wms.fis-wasser-mv.de` genutzt werden kann, um Wasserdaten aus dem Feature Type `app:eu_2021` zuverlässig abzurufen.

Dabei stehen die Attribute `eu_cd_gm` (im Sinne einer WFD-ID), die länderspezifische Messstellenummer, die Messgröße, das Probenahmedatum, Messwert und Einheit sowie die Bestimmungsgrenze (BG), der Indikator für die Unterschreitung der BG sowie der im Grundwasserbereich Trennkennzahl genannte Indikator für Gesamtprobe, filtrierte Probe oder Filterrückstand zur Verfügung.

Eine eingehendere Untersuchung der Daten förderte zutage, dass die mit dem Feld `eu_cd_gm` gelieferten WFD-IDs (Syntax: `DEGM_DEMV_00000000`) nur in vier Fällen mit der Syntax der WFD-ID in der derzeitigen Grundwasserdatenbank (Syntax meist: `DEGM_00000000`) übereinstimmt.

Dies ist für die Lokalisierung der Datensätze in den Beständen des UBA jedoch unschädlich, da dafür die Messstellenummer `mst_nr` mit sehr guter Übereinstimmung verwendet werden kann. Insofern ist eine Zuordnung der Messwerte aller Voraussicht nach im Produktiveinsatz möglich.

15.2 Ansatz zur medienübergreifenden Sicht und Recherche

Im Rahmen der Implementation des Fachinformationssystems wurde geprüft, in wie weit sich eine Volltextrecherche für die Fachanwendung eignet und welches Werkzeug gewählt werden sollte. Die Kriterien für die Auswahl waren die fachliche Eignung, die Möglichkeit einer kostenfreien Nutzung, die Möglichkeit der Anpassung durch Änderung des Quellcodes (OpenSource), Verfügbarkeit von Dokumentation sowie die Prognose einer möglichst langen Lebensdauer.

Hier ist untersucht worden, in wie weit z. B. Elasticsearch für eine integrierte und medienübergreifende Recherche eingesetzt werden kann. Elastic hat zwar im Jahr 2021 mit der Version 7.11 die Open-Source-Lizenzen durch die Server Side Public License ersetzt, die keine Open-Source-Lizenz ist, weil sie keine Cloudangebote erlaubt, jedoch ist eine Nutzung in der Cloud auch nicht Ziel des UBA.

Nachfolgend konnte durch eine Internetrecherche das beste OpenSource Searchindexer-Backend als Apache Lucene identifiziert werden. Das Search-Backend Lucene kann sowohl mit den Frontends Solr als auch mit Elastic betrieben werden. Wie erwähnt, stellt Elastic auf kommerzielle Services um und ist nur testweise kostenfrei, ansonsten ist der Vertrieb zu kontaktieren. Hingegen ist Solr weiterhin klar OpenSource und neben Lucene ebenfalls ein qualitätsgesichertes Projekt unter dem Dach der Apache Foundation. Damit erfüllen Solr und Lucene alle Anforderungen.

Folglich wurde versucht, die beiden Suchmaschinenteile Apache Lucene und Apache Solr in das FIS Wasser-Boden zu integrieren und die im FIS im Data Warehouse-Teil hinterlegten und hinterlegbaren Dokumente für Suchen zu indizieren und durchsuchbar zu machen. Dies konnte erfolgreich abgeschlossen werden und in der Folge steht der Implementation einer Volltextsuche auf dem Data Warehouse und den darin hinterlegten PDF-, Tabellen- und Präsentationsdokumenten im FIS Wasser-Boden aus IT-fachlicher Sicht nichts mehr im Wege.

Der für die Untersuchung der Eignung im FIS gewählte Standard-Tokenizer trennt an Leerzeichen, Tabulatoren, aber auch an Unterstrichen (_), Minus (-) und dem @. Er ist nicht dafür gedacht, Ziffernfolgen zu untersuchen. Dies kann mit dem Classic Tokenizer erreicht werden.

Die initiale Erstellung des Indexes benötigt wenige Sekunden Rechenzeit. Danach wird der real-time signal processor verwendet, um beim Hochladen neuer Dokumente und beim Ändern bestehender Dokumente genau diese neu zu indizieren. Bei Änderung eines Dokuments wird vor der Neugenerierung genau der Teil des Indexes entfernt, der zum alten Stand des geänderten Dokuments gehörte, so dass nicht versehentlich Dokumente gefunden werden, die den Text nicht mehr enthalten.

Solr/Lucene kennt mehr als 40 verschiedene Tokenizer und ist eines der am umfangreichsten konfigurierbaren Softwarewerkzeuge überhaupt. Sollte sich im Laufe der Nutzung des FIS Wasser-Boden die Möglichkeit ergeben, einen Search-Indexer produktiv in die Betriebsumgebung im Umweltbundesamt zu integrieren, kann ein passender Tokenizer gewählt und ggf. anders konfiguriert werden, als in diesem Forschungsvorhaben.

16 Ausblick

FIS Wasser-Boden ist bereits das erprobte Arbeitswerkzeug zur Berichterstattung an KOM und zur schnellen Recherche von Umweltdaten der damit abgedeckten Umweltkompartimente.

Eine sinnvolle, zukünftige Erweiterung des FIS Wasser-Boden, ist der Einsatz von künstlicher Intelligenz im erweiterten Sinn. Dies betrifft die Aufnahme von Bodendaten, für die bereits ein Training mit tausenden von Eingangsdatensätzen vorgenommen wurde. Die daraus gewonnenen Erkenntnisse sollten eingesetzt werden, um die Konvertierung von lieferant:innenspezifischen Listeneinträgen zu in FIS Wasser-Boden bekannten Listeneinträgen automatisiert vorzunehmen. Das ermöglicht perspektivisch die automatisierte Verarbeitung von Eingangsdaten und somit überhaupt die Aufnahme eines großen Teils von Bodendaten und spart Zeit und Arbeitskraft.

Weiterhin ist eine Integration der Meeresdatenbank MUDAB in das FIS Wasser-Boden geplant, sofern sich das Vorhaben organisatorisch begründen und finanzieren lässt. Damit würde sich der Umfang der über die Grenzen der Umweltkompartimente hinweg verfügbaren Umweltdaten deutlich erhöhen und die mit der Berichterstattung entstehenden, jährlichen Kosten voraussichtlich erheblich senken. Eine Beeinträchtigung der bestehenden Benutzererfahrung oder der Funktionalität entsteht dadurch aus unserer Sicht nicht.

Die Integration der Pflanzenschutzmitteldatenbank benötigt ebenfalls eine belastbare, organisatorische Begründung und Finanzierung. Aus technischer Sicht ist die Integration ebenso geboten, wie bei der MUDAB. Aufgrund zwar vorhandener Festlegungen zur Aufteilung der Arbeiten zwischen BMU und UBA, jedoch fehlender einschlägiger EU-Richtlinien, ist hier nicht absehbar, wann eine Integration erfolgen kann.

Aufgrund der bereits von Beginn an konsequent auf Erweiterbarkeit ausgelegten Struktur des Fachinformationssystems Wasser-Boden, aufgrund des Einsatzes der modernen OpenSource-Frameworks und -Datenbank, dem Bekenntnis zur IT-Architekturrichtlinie Bund sowie der Veröffentlichung aller Quellcodes²⁸ kann von einer langen Nutzungsperiode und zukünftigen, kostengünstigen Erweiterungen zur Unterstützung weiterer Prozesse und weiterer Umweltkompartimente fest ausgegangen werden.

²⁸ In den Quellcodes ist bei modernen Python/Django-Anwendungen das vollständige Datenmodell ebenfalls enthalten und zur Wiederverwendung zugänglich.

17 Quellenverzeichnis

- Aganina, K. (o. J.). *Digitalisierung in Deutschland: Was ist der aktuelle Stand?* Abgerufen 29. November 2023, von <https://www.ratbacher.de/blog/digitalisierung-stand-deutschland/>
- Bundesministerium des Innern und für Heimat (2017). *Gesetz zur Verbesserung des Onlinezugangs zu Verwaltungsleistungen (Onlinezugangsgesetz – OZG)*. Bundesgesetzblatt Teil I, 3122, 3138. <https://www.gesetze-im-internet.de/ozg/>
- Bundesministerium für Wirtschaft und Klimaschutz (o. J. - a). *Digitalisierungsindex*. Abgerufen 29. November 2023, von <https://www.de.digital/DIGITAL/Navigation/DE/Lagebild/Digitalisierungsindex/digitalisierungsindex.html>
- Bundesministerium für Wirtschaft und Klimaschutz (o. J. - b). *Digitalisierung. Den digitalen Wandel gestalten*. Abgerufen 29. November 2023, von <https://www.bmwk.de/Redaktion/DE/Dossier/digitalisierung.html>
- Bundesregierung (o. J.). *Digitalstrategie der Bundesregierung: Strategie für einen digitalen Aufbruch*. Abgerufen 29. November 2023, von <https://www.bundesregierung.de/breg-de/service/datenschutzhinweis/digitalstrategie-2072884>
- Codefirst (2017). *Tips for Managing Fixed Price Projects in Agile*. Abgerufen 8. März 2022, von <https://www.codefirst.co.uk/blog/fixed-price-agile-projects/>
- Diehl, A. (2023). *Agile Methoden Übersicht – Die wichtigsten agilen Arbeitsmethoden*. dno. Abgerufen 11. November 2024, von <https://digitaleneuordnung.de/blog/agile-methoden>
- Elmasri, R. & Navathe, S. (2016). *Fundamentals of database systems* (7th ed.). Pearson International. <https://elibrary.pearson.de/book/99.150005/9781292097626>
- Google LLC — Suche (o. J.). *Digitalisierung Deutschland*. Abgerufen 29. November 2023, von <https://www.google.de/search?q=Digitalisierung+Deutschland>
- Hunt et. al. (2001). *Manifesto for Agile Software Development*. Abgerufen 8. März 2022, von <https://agilemanifesto.org/>
- IT-P GmbH (2021). *Welches Pricing-Modell eignet sich für agile Projekte?* Abgerufen 8. März 2022, von <https://www.it-p.de/blog/welches-pricing-modell-eignet-sich-fuer-agile-projekte/>
- Krusche & Company GmbH (2020). *Is 'Fixed Price Agile' A Unicorn or Platypus?* Abgerufen 8. März 2022, von <https://kruschecompany.com/fixed-price-agile/>
- Leser, U. & Naumann, F. (2006). *Informationsintegration: Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen* (1. Auflage). dpunkt.verlag. <https://www.assets.dpunkt.de/openbooks/informationsintegration.pdf>
- Maggiori, E. (2023). *Tech-Branche: Wir arbeiten nicht. Null. - Golem.de*. Abgerufen 29. November 2023, von <https://www.golem.de/news/tech-branche-wir-arbeiten-nicht-null-2304-173002.html>
- Microsoft Corporation — Bing-Suche (o. J.). *Digitalisierung Deutschland*. Abgerufen 29. November 2023, von <https://www.bing.com/search?q=Digitalisierung+Deutschland>
- Moonkamp (2024). *Agile Methoden im Überblick: Scrum, OKRs, Kanban & Co*. Abgerufen 27. Juni 2024, von <https://mooncamp.com/de/blog/agile-methoden>
- Statista (o. J.). *Themenseite: Digitalisierung in Deutschland*. Abgerufen 29. November 2023, von <https://de.statista.com/themen/9036/digitalisierung-in-deutschland/>
- Umweltbundesamt (2013). *Kosten-Nutzen-Analyse*. Abgerufen 9. März 2022, von <https://www.umweltbundesamt.de/kosten-nutzen-analyse> (09.03.2022)

A Standardisierte Erhebungsstruktur zur Prozess- und Anforderungsanalyse

Die folgende standardisierte Erhebungsstruktur für eine Prozess- und Anforderungsanalyse ist für dieses Vorhaben entwickelt und auf jedes der in diesem Vorhaben betrachteten Umweltmedien separat angewandt worden. Sie dient als Hilfsmittel zur strukturierten Erhebung, auch wenn die Interviewpartner:innen viele Fragen nicht beantworten können. Die Vorlage stellt sicher, dass alle wesentlichen Aspekte angesprochen werden.

A.1 Status quo: Fachdatenbanken und Software

Was für eine Datenbank nutzt das derzeitige FIS?

- ▶ Gibt es mehrere Datenbanken?
- ▶ Gibt es Informationen, die außerhalb von Datenbanken gehalten werden?
- ▶ Wann wurden die Datenmodelle erarbeitet? Wurden diese regelmäßig gepflegt?

Welche Software wird zur Interaktion mit der Fachdatenbank eingesetzt?

- ▶ Gibt es Spezialsoftware? Welche Technologien werden dabei eingesetzt? Wann wurde diese erstellt? Wurde diese regelmäßig gepflegt?
- ▶ Welche Standardsoftware wird regelmäßig eingesetzt?
- ▶ Wird ggf. verschiedene Software für folgende Aufgabenbereiche genutzt?
 - Datenimport
 - Datenexport
 - Auswertungen
 - Berichterstellung

A.2 Datenarten

Welche Arten von Daten werden in der/den bestehende(n) Fachdatenbank(en) verarbeitet? Gibt es Datenarten, die derzeit noch nicht verarbeitet werden, aber für das neue FIS Wasser-Boden für dieses Umweltmedium berücksichtigt werden sollen?

Mögliche Datentypen sind nach Kategorie:

- ▶ Bewegungsdaten, z. B.:
 - Messdaten
- ▶ Stammdaten, z. B.:
 - Messstellen
 - geographische Entitäten (etwa Flüsse, Flusseinzugsgebiete, Grundwasserkörper)
 - administrative Entitäten (Messnetze)
- ▶ Referenzdaten, z. B.:

- Parameter, Messgrößen
 - physikalische Einheiten
 - Grenzwerte für Messwerte (etwa für Güteklassen oder zur Plausibilitätsbestimmung)
- Metadaten, z. B.:
- Quelle der Daten
 - Qualität der Daten

A.3 Datenmenge

- Wie umfangreich sind die Datenbestände?
- Um welchen Umfang wachsen die Datenbestände regelmäßig?
- Werden wesentliche Änderungen des Umfangs der Datenbestände erwartet?

A.4 Datenimport

Für jeden in A.2 „Datenarten“ bestimmten Datentyp der Kategorien Bewegungs-, Stamm- und Referenzdaten sollen folgende Fragestellungen geprüft werden:

A.4.1 Quelle

Aktualisieren sich diese Daten? Gibt es einen Prozess zum regelmäßigen Bezug der Daten?

Woher kommen die Daten?

- Werden die Daten durch Dritte geliefert? Wenn ja, wer sind die Lieferanten?
- Werden die Daten extrahiert? Wenn ja, woher?

Wie zuverlässig ist dieser Prozess?

- Gibt es eine (ggf. gesetzliche) Grundlage, die die Zurverfügungstellung oder Lieferung der Daten garantiert?
- Sind Änderungen an dieser Grundlage geplant?
- Hat das UBA die Möglichkeit, Änderungen an der Grundlage vorzunehmen oder darauf einzuwirken?

Gibt es ggf. verschiedene Quellen? (Bei Abweichung zwischen verschiedenen Quellen, sollten die folgenden Punkte je Quelle diskutiert werden.)

A.4.2 Übermittlungsmodus

Wie werden die Datensätze derzeit übermittelt?

- Werden die Daten auf Anforderung geliefert bzw. geholt oder gibt es einen Turnus?

- ▶ Wie werden die Daten übermittelt? (etwa per E-Mail, per Upload, per Download)
- ▶ In welchem Format werden die Daten übermittelt?
- ▶ Welchen Umfang haben die übermittelte Datensätze? Ist dieser stets gleich? Wird stets der volle Umfang übermittelt oder kann die Übermittlung in mehreren Teilen ggf. zu verschiedenen Zeitpunkten stattfinden?

Gibt es geplante oder gewünschte Änderungen des Übermittlungsmodus?

A.4.3 Import in die Fachdatenbank

Wie werden die Daten derzeit in die Fachdatenbank übertragen?

- ▶ Gibt es eine automatische Importroutine im FIS oder werden die Daten manuell übertragen? Mit welchen Werkzeugen?
- ▶ Müssen die Daten dabei transformiert werden?
- ▶ Wird stets der gesamte Datensatz importiert oder können auch nur Teile importiert werden?

A.4.4 Qualitätssicherung

Werden die Daten geprüft? Wenn ja,

- ▶ wann: Vor dem Import in die Fachdatenbank oder zu einem späteren Zeitpunkt in der Fachdatenbank?
- ▶ wie: manuell oder automatisch? Mit welchen Werkzeugen?
- ▶ was: syntaktisch, semantisch, inhaltlich?

Wie ist das Vorgehen bei der Beanstandung und Korrektur der Daten?

- ▶ Werden Fehler selbst korrigiert? Wenn ja, welche Fehlertypen betrifft das und wie wird dabei vorgegangen? Wird das dokumentiert, ggf. wie?
- ▶ Werden Fehler den Datenoriginator:innen gemeldet und von diesen korrigiert?

Wie ist das Vorgehen beim Import von fehlerhaften Datensätzen in die Fachdatenbank?

- ▶ Werden fehlerhafte Datensätze vollständig importiert, wird nur der fehlerfreie Teil importiert oder wird der ganze Datensatz zurückgewiesen?
- ▶ Bei vollständigem Import der Daten: Werden fehlerhafte Daten gekennzeichnet oder nach dem Import wieder gelöscht?

Gibt es gewünschte Änderungen?

- ▶ Gibt es QS-Prüfungen, die sich als unnötig oder unpraktikabel erwiesen haben?
- ▶ Gibt es Bedarf für zusätzliche QS-Prüfungen?

A.4.5 Datenkorrektur und -aktualisierung

Bei Bewegungsdaten: Wie werden Korrekturen von bereits in die Fachdatenbank importierten Datensätzen importiert?

- ▶ Beinhaltet der Korrekturdatensatz auch die bereits fehlerfreien Daten des Originaldatensatzes oder enthält er nur die Korrekturen der fehlerhaften Daten?
- ▶ Werden die Daten überschrieben? Wenn ja: Wie werden diese identifiziert?
- ▶ Werden die zuvor importierten Daten vor dem Korrekturimport gelöscht?

Bei Stamm- und Referenzdaten: Werden Änderungen bestehender Datensätze berichtet? Wie wird dort beim Import vorgegangen?

- ▶ Werden Stilllegungen erfasst?
- ▶ Werden Vorgänger-Nachfolger-Relationen erfasst?
- ▶ Werden zeitliche Änderungen von Eigenschaften erfasst?

Besteht Bedarf, hierbei von der bisherigen Vorgehensweise abzuweichen?

A.5 Manuelle Datenänderungen

Können Daten in der Fachdatenbank manuell (also nicht im Rahmen des Datenimports) erzeugt, geändert oder gelöscht werden? Wenn ja:

- ▶ Welche Daten betrifft das?
- ▶ Wird diese Funktionalität genutzt?
- ▶ Woraus resultiert dieser Änderungsbedarf?

Werden die manuellen Änderungen dokumentiert? Können diese rückgängig gemacht werden?

Besteht Bedarf, hierbei von der bisherigen Vorgehensweise abzuweichen, neue Möglichkeiten zu schaffen oder ist die bestehende Funktionalität teilweise überflüssig?

A.6 Datenanalyse

A.6.1 Methodik

Welche Auswertungen werden regelmäßig durchgeführt? Gibt es häufig spezielle, einmalige Auswertungen oder ändert sich der Auswertungsbedarf häufig?

Wie werden regelmäßige und unregelmäßige Auswertungen vorgenommen? Manuell oder durch das FIS unterstützt? Welche Werkzeuge kommen zum Einsatz?

Werden Daten regelmäßig graphisch visualisiert? Wenn ja, wer macht dies und welche Werkzeuge kommen zum Einsatz?

A.6.2 Berechnungen

Welche aufwändigen Berechnungen sind regelmäßig nötig?

- ▶ zeitliche Aggregationen?
- ▶ geographische Aggregationen?
- ▶ Aggregationen über Parameter? (z.B. Ionenbilanzen, Nitrat und Nitrat-N, o. ä.)?
- ▶ Klassierung von Messwerten? Häufigkeitsverteilungen?

Wann finden diese Berechnungen derzeit statt?

- ▶ Bei der Auswertung
- ▶ Bereits beim Datenimport (Ergebnisse werden in der Datenbank gespeichert)
- ▶ Manuell zu jedem beliebigen Zeitpunkt (Ergebnisse werden in der Datenbank gespeichert)

A.7 Datenexport

Gibt es Bedarf für manuellen Export von Rohdaten zur Weiterverarbeitung außerhalb des FIS?
Wenn ja,

- ▶ In welchem Format? Datenbankdump (SQL), Excel, CSV?
- ▶ Welche Datenarten werden exportiert?
- ▶ Was sind typische Anwendungsfälle?
- ▶ Gibt es weitere Anforderungen (etwa Filter)?

A.8 Berichterstattung

Welche Berichtsprozesse gibt es derzeit oder wird es künftig geben? Für jeden Berichtsprozess stellen sich folgende Fragen:

- ▶ Wer ist Empfänger des Berichts?
- ▶ In welchem Turnus wird berichtet? Oder wird auf Anforderung berichtet?
- ▶ Auf welcher gesetzlichen Grundlage wird berichtet?
- ▶ In welchem Format wird bisher berichtet? In welchem Format soll/darf berichtet werden?
- ▶ Wie werden die Daten bisher übermittelt? Wie sollen/dürfen die Daten übermittelt werden?
- ▶ Sind für die Erstellung des Berichts externe Daten oder Informationen nötig, die nicht Teil des FIS sind?
- ▶ Wie wird bei Nachforderungen und Korrekturen vorgegangen?

A.9 Verschiedenes

- ▶ Welche Funktionen des FIS wurden in den Vergangenheit nicht oder extrem selten gebraucht?
- ▶ Welche im FIS verwaltete Daten wurden in den Vergangenheit nicht oder extrem selten gebraucht?
- ▶ Welche Funktionen oder Daten fehlten in der Vergangenheit?

B Prozessanalyse Fließgewässer – Ergebnisse

Die im folgenden Text festgehaltenen Aussagen sind in die standardisierte Erhebungsstruktur (siehe Anhang A) eingebettet, die für die Konzeptentwicklung erarbeitet wurde. Der Text enthält die Ergebnisse aus den Interviews und den Analysen der übersandten Ressourcen (Datenbankabzüge, Beispieldateien etc.).

B.1 Status quo: Fachdatenbanken und Software

Das FIS Fließgewässer „QUADAWA“ nutzt eine SQLServer-Datenbank. Die Datenbank und die zugrundeliegenden Datenmodelle wurden 2005 erstellt und bei Bedarf angepasst.

Als Benutzerfrontend wird Microsoft Access eingesetzt. Für den Import von ASCII/TXT-Dateien und die Auswertung von Daten wurde ein Frontend in MS Access programmiert. Davon werden aber in der Regel nur die Datenimport-unterstützenden Funktionen verwendet.

Mithilfe von Microsoft Access können selektierte Daten außerdem im .xlsx-Format exportiert werden. Zur weiteren Auswertung wird dann Microsoft Excel genutzt.

B.2 Datenarten

B.2.1 Stammdaten

► Messstellen (QUADAWA-Tabelle: station)

- Messstellen haben eine vom FG II 2.4 und eine von den Lieferant:innen vergebene Identifikationsnummer.
- Messstellen sind entweder dem Bundesland, in dem sie liegen oder einer der vier Kategorien „Bund“, „E-B-Institut“, „GKSS“, „HYDABA BfG Berlin“ zugeordnet. Dies bezeichnet den/die Datenlieferant:innen. Der/die Datenlieferant:in wird benötigt.
- Knapp zwei Drittel der Messstellen besitzen eine Angabe des Stromkilometers und gut die Hälfte besitzen eine Angabe zur Lage im Fluss. Weiterhin sind Angaben zur Blattnummer der topografischen Übersichtskarte (topo_karte), zu Rechts- und Hochwert, zur Größe des Flusseinzugsgebiets, zum Institut sowie Höhenlage und LAWA-Kennnummer möglich und bei zahlreichen Datensätzen auch vorhanden. Die Angaben entsprechen den Vorgaben des Anhangs der VwV.
- Messstellen können stillgelegt werden, das wird am Datensatz aber nicht gesondert markiert, sondern kann nur dadurch erkannt werden, dass für sie keine Messwerte mehr aufgenommen werden.

► Messnetze

- Messstellen werden zu verschiedenen administrativen Zwecken zu Messnetzen gruppiert. Messstellen können Teil mehrerer oder keines Messnetzes sein.
- Die Zugehörigkeit von Messstellen zu Messnetzen wird in der QUADAWA durch die Spalte lawa_kennnr angezeigt.

► Wasserkörper

- Die Zuordnung von Messstellen zu Wasserkörpern ist für einige Auswertungen (z. B. EUA-Bericht) nötig. Die Angaben der WRRL-Berichterstattung (BfG-Daten: Wasserkörper, Station-Code nach WRRL) werden für die Meldung nach WISE genutzt. Die Pflege dieser Angaben bei der BfG erfolgt durch die Länder. Es ist ein halbjährliches Update vorgesehen. Daher werden die Angaben nicht in `station` gespeichert, sondern aktuell abgefragt und über die Angabe EIONET verknüpft und bei Erstellung der Exportdateien berücksichtigt.

B.2.2 Referenzdaten

- ▶ Einheiten (QUADAWA-Tabelle: `einheit`): Die Codes sind ein Auszug der LANUV-Liste. Es gibt keine UBA-interne Codierung. Codes werden ggf. manuell ergänzt.
- ▶ Einheitenumrechnungsvorschriften (QUADAWA-Tabelle: `eiumrech`): DB-interne Angaben, um beim Import die Daten ggf. in die für die DB festgelegte Einheit umzurechnen. Angaben können ggf. manuell ergänzt werden.
- ▶ Stoffe (`stoff`): Angaben entsprechen i.d.R. der LANUV-Stoffliste. Derzeit sind noch DB-interne Codes enthalten. Im FIS sollen nur die LANUV Codes enthalten sein. Die Tabelle wird vor endgültiger Übergabe ins FIS bearbeitet. Die Tabellenbeschreibung wurde übergeben. QS: Ein Stoff kann nur gestrichen werden, wenn keine Verknüpfungen zu `paranr` und Bewegungsdaten bestehen.
- ▶ Parameter (`paranr`): Die Tabelle enthält die Matrix und die Einheit der Stoffe. Die Datei ist unbedingt erforderlich und wird manuell fortgeschrieben. QS: Ein Eintrag ist nur möglich, wenn der Stoffcode (`parameter4`) in `stoff` enthalten ist.
- ▶ Stoffgruppen (QUADAWA-Tabellen: `gruppe` und `ogruppe`): Interne Tabellen für die Abfrage von Stoffgruppen.
- ▶ Probenahmematrizen (Spalte `parameter5` der QUADAWA-Tabelle `einwerte` und `aggwerte`): Siehe Tabelle `paranr`
- ▶ Probearten (QUADAWA-Tabelle `probeart`): Siehe Tabellenbeschreibung
- ▶ Messarten (Spalte `messart` der QUADAWA-Tabelle `einwerte` und `aggwerte`): s. Tabellenbeschreibung
- ▶ Güteklassen (Wertebereiche (je Parameter) für die Klassifizierung von Messwerten zur Auswertung der Wasserqualität): Interne Tabelle für Auswertung pro Güteklassen, manuelle Eingabe, Fortschreibung bei Bedarf

B.2.3 Bewegungsdaten

- ▶ Messwerte:
 - Messwerte (Tabelle `einwerte`) werden je Station, Parameter, Matrix, Probeart (`akz`), Messzeit und Messart erfasst.
 - Neben dem Messwert wird die Bestimmungsgrenze (Pflichtfeld) erfasst.

- Mit einem Kennzeichen kann angegeben werden, dass der gemessene Wert unter ('-') oder über ('+') der Bestimmungsgrenze liegt. In diesem Fall wird nur die Bestimmungsgrenze, nicht aber der Messwert erfasst.
- Messwerte und Bestimmungsgrenze werden in der dem Parameter zugeordneten Einheit gespeichert.
- Bei Mischproben und kontinuierlichen Messungen gibt das Feld Messprogrammwoche die Länge des Zeitraums in Tagen an, in dem die Mischprobe genommen wurde. Die Messzeit gibt einheitlich den Anfang der Periode an.

► aggregierte Messwerte

- Aggregierte Werte werden analog zu Einzelwerten je Station, Parameter, Matrix, Probeart (akz), Messzeit und Messart erfasst.
- Sie enthalten das Aggregationsintervall (vrs. in Tagen, Spalte probeint), die Anzahl der Einzelwerte, die Anzahl der Einzelwerte unterhalb der Bestimmungsgrenze (Spalte anzwertubg), die Bestimmungsgrenze, numerische Werte zur deskriptiven Statistik (minimum, perz_10, perz_50, mittelwert, perz_90, maximum), Textfelder zur deskriptiven Statistik (mini, p10, p50, mw, p90, maxi), die bei Unterschreitung der Bestimmungsgrenze die Zeichen '<' vorangestellt haben und die bei geringer Anzahl der zugrundeliegenden Einzelwerte für die Perzentile und den Mittelwert die Zeichenkette '*****' enthalten.
 - Minimum und Maximum werden immer ausgewiesen
 - Mit zunehmender Zahl der Einzelwerte, werden zunehmend auch Mittelwert (ab 3 Messwerten), Median, 90 %- (ab 11 Messwerten) und 10 %-Perzentil (ab 16 Messwerten) ausgewiesen (Abweichung bei Altdaten)

B.2.4 Metadaten

Der Importzeitpunkt von Messwerten bzw. der Berechnungszeitpunkt von aggregierten Messwerten wird gespeichert.

Weiterhin besitzen die Messwerte (Tabelle einwerte und aggwerte) einen Indikator über die Validierung für die Freigabe bei Datenrecherchen. Der Indikator wird bisher nicht genutzt, da nur FG II 2.4 Zugriff auf die Daten hat.

B.3 Datenmenge

B.3.1 Stammdaten

- 604 Messstellen (davon ca. 300 derzeit aktive, also nicht stillgelegte Messstellen), Messstellen "E-B-Institut", "GKSS" werden nicht ins FIS übergehen.
- 3 Messnetze (LAWA, Watchlist, nationale Beobachtungsliste), teils Untermengen, teils Überschneidungen: interne Zuordnung für Exportfunktionen und Auswertungen im FG II 2.4.

B.3.2 Referenzdaten

- ▶ 24 Einheiten (Tabelle `einheit`)
- ▶ 1278 Stoffe (Tabelle `stoff`, Eindeutigkeit durch Spalte `parameter4` (LAWA-Nummer))
- ▶ 1873 Parameter (Tabelle `paranr`, Eindeutigkeit durch die Kombination der Spalten `parameter4` und `parameter5` (Matrix))
- ▶ 6 Probenahmematrizen: gelöst (6), Gesamtprobe (1), abgesetzt, Überstand für Analyse (3), Schwebstoff/Sediment (0), Fische (8), Muscheln (9); Angaben müssen erweiterbar sein
- ▶ 5 Probearten (B - Biota-Fischfilet, F - Filterrückstand, S - Absetzbecken, W - Wasserphase, Z - Zentrifuge)
 - Es gibt bisher keine Messwerte zu zentrifugierten Proben. Messwerte sind überwiegend der Wasserphase entnommen.
- ▶ 3 Messarten (s. Anhang VwV und Dateibeschreibung)
 - „E“: Einzelprobe (7,2 Millionen Werte)
 - „M“: Mischprobe (1 Million Werte)
 - „K“: kontinuierliche Messung (0,3 Millionen Werte)
- ▶ 7 Güteklassen für 53 Stoffe (Messgrößen) mit Angabe, ob der Vergleich mit kleiner oder größer Klassengrenze erfolgen muss. (Spalte: `kleiner`)

B.3.3 Bewegungsdaten

- ▶ 8,6 Millionen Messwerte (Tabelle `einwerte`)
- ▶ 1,1 Millionen aggregierte Werte (Tabelle `aggwerte`)

B.4 Datenimport

B.4.1 Datenimport Messstellen

Quelle

Bis auf die im Folgenden genannten Ausnahmen werden alle Attribute von den Landesumweltbehörden gemeldet.

Die Messstellen-ID (QUADAWA-Spalte: `stat_lawa`) ist Primärschlüssel für Messstellen, wird vom FG II 2.4 vergeben und den Landesumweltbehörden zurückgemeldet. Die ID beginnt immer mit dem Kürzel des Bundeslands. Handelt es sich bei einer neuen Messstelle um den Ersatz einer bisherigen Messstelle, die geographisch sehr nah liegt, vergibt das FG eine neue ID bestehend aus der originalen ID ergänzt um eine fortlaufende Ziffer. Bei Auswertungen über zeitliche Verläufe von Messwerten werden diese Messstellen in der Regel als identisch betrachtet.

Die Länder vergeben unterschiedliche Messstellennummern pro Matrix für dieselbe Messstelle. Der Wasserblick übernimmt deren div. Nummern. Die EUA möchte (verständlicherweise) nicht unterschiedliche Nummern, sondern nur genau eine.

Die WFD-Nummer (auch WRRL-Nummer oder Wasserblick-Nummer) von Messstellen wird beim Wasserblick erfragt. Die Zuordnung der WFD-Nummern zu den Messstellen in der Datenbank erfolgt über die EIONET-Nummer, die Pflege im Wasserblick durch die Länder erfolgt derzeit nur im Rahmen des Reportings zur WRRL. Dies soll auf ein halbjährliches Intervall geändert werden. Bei der Erstellung des Exportmoduls im FIS muss eine Verknüpfung mit der Tabelle SWSTN aus dem WasserBLiCK vorgesehen werden.

Übermittlungsmodus

Die Verwaltungsvereinbarung hat eine ASCII-Vorlage für die Stammdatenübermittlung. Wenn es sich um einzelne Messstellen handelt, wird diese jedoch von den Datenlieferant:innen nicht genutzt. Die Angaben werden per E-Mail übermittelt.

Im FIS ist vorzusehen, dass station vom UBA gepflegt wird. Eine Pflege durch die Länder ist nicht praktikabel.

Import in die Fachdatenbank

Die Daten werden manuell mithilfe von Microsoft Access bzw. von Modulen der Datenbankanwendung in die Datenbank eingetragen.

Qualitätssicherung

entfällt

Datenkorrektur und -aktualisierung

Bei Korrekturen oder Aktualisierungen wird der bestehende Datensatz manuell überschrieben. Eine Versionierung findet nicht statt. Der überschriebene Stand ist nicht mehr verfügbar. Das Erfassen des Datums der letzten Änderung ist aus fachlicher Sicht nicht erforderlich.

Es gibt seitens FG II 2.4 keinen Änderungsbedarf an diesem Vorgehen.

B.4.2 Messnetze

Quelle

Festlegung durch UBA/II 2.4, Anzahl variabel

Übermittlungsmodus

nicht erforderlich

Import in die Fachdatenbank

manuelle Eingabe

Qualitätssicherung

nicht erforderlich

Datenkorrektur und -aktualisierung

Unregelmäßig; bei Bedarf

B.4.3 Einheiten

Quelle

Auszug aus dem Einheitenkatalog des LANUV NRW

Übermittlungsmodus

nicht erforderlich

Import in die Fachdatenbank

manuell

Qualitätssicherung

nicht erforderlich

Datenkorrektur und -aktualisierung

Manuell, bei Bedarf

B.4.4 Einheitenumrechnungsvorschriften

Die Einheitenumrechnungsvorschriften sind statisch. Der Datenbestand ist bekannt und ändert sich selten.

B.4.5 Datenimport Stoffe/Parameter

Quelle

Derzeit:

Die Liste entspricht einem Teilausschnitt des Stoffkatalogs des LANUV NRW und wird durch das FG II 2.4 durch zusätzliche, im LANUV-Katalog fehlende Stoffe ergänzt.

Die Stoffe sind durch eine alphanumerische ID kodiert. Stoffe aus dem Katalog des LANUV sind numerisch, die vom FG eingetragenen Stoffe erhalten vom FG eine, nicht mit dem Namensschema der LANUV-Stoff-IDs kollidierende, textuelle Stoff-ID. Wird der Stoff später in den LANUV-Katalog aufgenommen, wird die LANUV-Stoff-ID übernommen.

Planung für das FIS:

Im FIS werden nur noch die LANUV-Stoff-IDs geführt. FG-Codes werden nicht mehr genutzt.

Problem:

Der Ursprung der Stoffnamen ist nicht unerheblich, da Stoffe in Grundwasser teilweise anders benannt sind, als in Fließgewässern. Es gibt alternative Namen und jedes FG hat seine bevorzugten Namen. Die Stofftabelle Fließgewässer darf nicht geändert werden in dem Sinne, dass das FIS Wasser-Boden für den Bereich Fließgewässer weiterhin die dort bekannten Stoffnamen (OGewV oder LANUV-Namen) – und nur diese – verwenden können muss und die LANUV-Stoff-ID ein alphanumerisches Feld bleiben muss.

Die Tabellen Stoffe und Parameter werden unregelmäßig aktualisiert, etwa wenn der LANUV-Katalog aktualisiert wird und das FG Änderungen übernimmt oder wenn das FG fehlende Stoffe erkennt und ergänzt.

Es werden etwa 20 neue Stoffe pro Jahr aufgenommen.

Übermittlungsmodus

Per E-Mail vom LANUV, auch über WasserBLiCK verfügbar

Import in die Fachdatenbank

Manuell, da auch eine Zuordnung zu den Stoffgruppen erforderlich ist

Qualitätssicherung

Die Prüfung erfolgt durch die Verknüpfung der Dateien, dadurch wird geprüft, ob ein Datensatz vollständig ist oder gelöscht werden kann; weitere QS-Prüfungen entfallen.

Datenkorrektur und -aktualisierung

Die Stoffliste wird manuell korrigiert. Eine Versionierung findet nicht statt. Der überschriebene Stand ist nicht mehr verfügbar.

B.4.6 Datenimport Stoffgruppen

Interne Tabelle, weitgehend statisch, manuelle Eingaben ohne Qualitätssicherung

B.4.7 Datenimport Probenahmematrizen

Probenahmematrizen sind weitgehend statisch in dem Sinne, dass bestehende Matrizes nicht entfallen. Es müssen jedoch zukünftig auch neue Matrizes ergänzt werden können.

B.4.8 Datenimport Probenahmearten

Probenahmearten sind statisch. Der Datenbestand ist bekannt und ändert sich kaum. Ergänzungen müssen möglich sein.

B.4.9 Datenimport Messarten

Messarten sind statisch. Der Datenbestand ist bekannt und ändert sich kaum. Ergänzungen müssen möglich sein.

B.4.10 Datenimport Güteklassen

statisch. Der Datenbestand ist bekannt und ändert sich kaum. Ergänzungen müssen möglich sein.

B.4.11 Datenimport Messwerte

Quelle

Messwerte werden von den 16 Landesumweltbehörden und der BfG (für Messwerte zweier Messstellen aus Rheinland-Pfalz) berichtet.

Die rechtliche Grundlage für die Übermittlung der Messwerte ist der Anhang 1 der VwV zum Datenaustausch zwischen Bund und Ländern.

Übermittlungsmodus

Die Übermittlung erfolgt per E-Mail.

Die Übermittlung aller Messwerte eines Kalenderjahres erfolgt jährlich zum 31.07. des Folgejahres. Die Fristen werden nicht immer eingehalten.

Je Probematrix wird eine Datei mit den Messwerten des Kalenderjahres gesendet. Die Probematrix kann aus dem Dateinamen abgelesen oder aus Begleitinformationen in der E-Mail extrahiert werden.

Die Lieferungen kommen mitunter in mehreren Teillieferungen, insbesondere, wenn die Frist nicht eingehalten werden kann.

Es gibt eine vorgegebene Tabellenstruktur. Die Dateiformate variieren zwischen CSV, XLSX und ASCII (Festbreiten-TXT). Die Tabellenstruktur wird nicht immer eingehalten, insbesondere der Spaltenumfang, die Spaltenkopfbeschreibungen und die Spaltenreihenfolge kann abweichen. Auch innerhalb einer Lieferung mit mehreren Dateien stimmt das Format nicht immer zwischen allen Dateien überein.

Import in die Fachdatenbank

Der Import erfolgt immer über das Frontend als ASCII-Datei.

Zur Erstellung der Importdatei kann im FG eine ACCESS-Anwendung genutzt werden, in der die EXCEL-Dateien in die ASCII-Datei exportiert werden.

Qualitätssicherung

Das Importmodul prüft automatisch bei allen Messwerten, ob alle Pflichtangaben vorhanden sind und ob Messstelle und der Stoff der QUADAWA bekannt sind. Weicht die Einheit von der in der QUADAWA für den Stoff vorgesehenen Einheit ab, wird der Messwert automatisch umgerechnet.

Datensätze, die von der QS erkannt werden, werden in eine Fehlerdatei mit einem zusätzlichen Fehlerkommentar eingespielt. Die Datei wird an die Länder zur Prüfung und Korrektur übermittelt.

Das Prüfungsergebnis wird dem UBA per E-Mail übermittelt und in das Frontend übernommen. Danach erfolgt ein erneutes Einspielen über das Frontend in die Datenbank. Bei größeren Datenkorrekturen werden alle Messwerte für ein Jahr und Land gelöscht und neu importiert. Dieses Verfahren ist in der IT üblich und soll bei den Fließgewässern weiterhin Bestand haben. Ggf. soll auch die Löschung auf einen bestimmten Stoff/Messstelle limitiert werden können.

Eine manuelle Anpassung von Messwerten vor dem Import erfolgt nicht.

B.5 Datenänderungen

Eine Aktualisierung der Stofflisten und der Messstellen kann über die bestehende Anwendungsoberfläche vorgenommen werden. Diese Funktionalität wird jedoch zugunsten manueller Bearbeitung auf der Datenbank selten genutzt.

Alle Daten können manuell auf der Datenbank geändert werden. Eine Dokumentation oder Versionierung ist derzeit nicht vorgesehen.

Die Möglichkeit der manuellen Änderung wird zukünftig eher nicht notwendig sein, wenn es eine Benutzungsoberfläche für alle Tabellen gibt, bei denen manuelle Änderungen oder Ergänzungen notwendig sind.

B.5.1 Aggregierte Messwerte

Aggregierte Messwerte werden aus der Messwerttabelle berechnet (Ausnahme: Daten vor 2008).

B.6 Datenanalyse

B.6.1 Methodik

Die Auswertungsroutinen entstammen dem Jahr 2005 und sind in Access abgelegt. Sie werden in der Regel nicht genutzt, weil FG II 2.4 Auswertungen durch Abfragen in MS Access schneller und flexibler selbst durchführen kann.

Mithilfe einer Exportfunktion in das XLSX-Format werden Auswertungen und graphische Visualisierungen (z.B. Balkendiagramme) teilweise auch mit MS Excel vorgenommen.

Übliche Auswertungen sind die Prüfung auf Ausreißer, Vergleich mit Vorgaben (z.B. Umweltqualitätsnormen) und Zuordnung zu Güteklassen.

B.6.2 Berechnungen

Die Tabelle der aggregierten Messwerte wird mit einer SQL-Routine aus der Messwerttabelle berechnet. Diese Berechnung wird manuell ausgelöst, weil sie sehr zeitaufwändig ist. Die Berechnung kann zeitlich, stofflich und nach Messstellen eingeschränkt werden.

B.7 Datenexport

Die Daten werden derzeit in EXCEL-Dateien exportiert nach den Vorgaben der:

- ▶ EUA (WISE 6),
- ▶ Watch List Durchführungsbestimmung (WISE 6, aber andere Messstellenauswahl)
- ▶ Nitratrichtlinie

Weitere flexible Exporte nach EXCEL werden über MS ACCESS für interne Auswertungen erstellt.

B.8 Berichterstattung

Das UBA muss regelmäßig einen Nitratbericht und einen EUA-Bericht aus der QUADAWA generieren. Diese werden vom FG mittels eigener Abfragen im Access im Excel-Format generiert.

Für den EUA-Bericht werden derzeit aggregierte Werte berichtet, ab 2022 müssen aber Einzelwerte berichtet werden.

B.9 Verschiedenes

Die Funktionalität von QUADAWA soll im neuen FIS Wasser-Boden erhalten bleiben, das ist FG II 2.4 von den Projektleitern zugesichert worden. Das heißt, dass mit dem MS Access-

Programm wie bisher lesend auf die Daten in der bisherigen Struktur zugegriffen werden können soll, sodass Auswertungen und Berichterstellungen weiterhin wie gewohnt durchgeführt werden können. Die Prozesse zum Import und der Änderung bzw. Korrektur von Daten dürfen überarbeitet werden.

B.9.1 Unnötige Funktionen oder Daten

Es sind derzeit keine unnötigen Funktionen oder Daten bekannt.

B.9.2 Fehlende Funktionen oder Daten

Es fehlt die Möglichkeit, anzugeben und zu speichern, welche Fischart bzw. welche Muschelart untersucht wurde. Dies muss auch in der VwV geregelt werden.

C Prozessanalyse Boden

Die im folgenden Text festgehaltenen Aussagen sind in die standardisierte Erhebungsstruktur eingebettet, die für die Konzeptentwicklung erarbeitet wurde und enthält die Ergebnisse aus den Interviews und den Analysen der übersandten Ressourcen (Datenbankabzüge, Beispieldateien etc.).

C.1 Status quo: Fachdatenbanken und Software

Im Umweltkompartiment Boden wird die Datenbank eBIS mit der gleichnamigen Webanwendung eingesetzt. eBIS enthält auch die Daten des F+E-Vorhabens „Integration von im UBA vorhandenen Bodendaten in eine Fachanwendung zur bundesweiten Auswertung“, FKZ 3717 74 252 0. Perspektivisch sollen Daten aus dem Internet (z. B. Klimadaten und Verdichtungsräume) ebenfalls genutzt werden.

Das Datenmodell von eBIS wurde im Jahr 2007 von UBA FG II 2.7 (H. Böken) in Zusammenarbeit mit Firma RISA (M. Lüttgert) entwickelt und in den Jahren 2009 bis 2011 von UBA FG II 2.7 und Firma ENDA wesentlich überarbeitet. Nachfolgend wurde das Datenmodell im Rahmen von Pflegearbeiten noch graduell geändert.

Neben der Webanwendung eBIS wird im UBA direkt auf der Postgres-Datenbank lesend gearbeitet. Dazu werden die Anwendungen pgAdmin 4, DBeaver und DbVisualizer eingesetzt. Neben einer Oberfläche zur Anzeige und Bearbeitung von Datensätzen verfügt eBIS als „Spezialsoftware“ über ein Abfragewerkzeug, ein Abfragewerkzeug speziell für Boden-Pflanze-Wertepaare sowie ein Mapping-Tool. Das Mapping-Tool ermöglicht es, im UBA bzw. in eBIS unbekannte Listeneinträge, die aus den Datenlieferungen der Bundesländer stammen, bekannten Listeneinträgen aus eBIS zuzuordnen. Diese Zuordnungsinformationen werden zukünftig für die schnellere Übernahme von Länderdatenlieferungen ausgewertet.

eBIS basiert auf der Entwicklungsplattform Python/Django für Webanwendungen und wird regelmäßig gepflegt. Neben eBIS werden folgende Softwareprodukte genutzt: QGIS, ArcGIS, R und Excel.

C.2 Datenarten

Welche Arten von Daten werden in der/den bestehende(n) Fachdatenbank(en) verarbeitet? Gibt es Datenarten, die derzeit noch nicht verarbeitet werden, aber für das neue FIS Wasser & Boden für dieses Umweltmedium berücksichtigt werden sollen?

Mögliche Datentypen sind nach Kategorie:

- ▶ Bewegungsdaten, z. B.:
 - Messdaten: Werden bereits in eBIS gehalten und verarbeitet.
- ▶ Stammdaten, z. B.:
 - Messstellen: Werden bereits in eBIS gehalten und verarbeitet.
 - geographische Entitäten (etwa Flüsse, Flusseinzugsgebiete, Grundwasserkörper): Derzeit sind Naturräume als Liste in den Untersuchungsflächen referenziert. Es ist jedoch geplant, diese und Informationen zu Klimastationen und Klima, Verdichtungsräume und

administrative Entitäten wie Gemeinden auf Basis von Karten oder Kartendiensten einzubinden.

- administrative Entitäten (Messnetze): Bundesländer und (s.o.) Verdichtungsräume werden bereits in eBIS gehalten und verarbeitet.

► Referenzdaten, z.B.:

- Parameter, Messgrößen: Werden bereits in eBIS gehalten und verarbeitet.
- physikalische Einheiten: Werden bereits in eBIS gehalten und verarbeitet.
- Grenzwerte für Messwerte (etwa für Güteklassen oder zur Plausibilitätsbestimmung): Werden derzeit NICHT in eBIS gehalten.

► Metadaten, z.B.:

- Quelle der Daten: Werden bereits in eBIS gehalten und verarbeitet.
- Qualität der Daten: Werden bereits in eBIS gehalten und verarbeitet (Messunsicherheit, Konfidenz algorithmisch, Konfidenz Expertin, Konfidenz gesamt). Die Konfidenzen werden durch fünf Vertrauensklassen abgebildet.

C.3 Datenmenge

► Wie umfangreich sind die Datenbestände?

- eBIS enthält derzeit 2,7 Millionen Bodenmesswerte und 0,47 Millionen Pflanzenmesswerte aus 87 Datenlieferungen.

► Um welchen Umfang wachsen die Datenbestände regelmäßig?

- Qualitative Antwort: gering (Daten aus F&Es und 4-5jährige Lieferungen aus einigen Bundesländern)

► Werden wesentlichen Änderung des Umfangs der Datenbestände erwartet?

- Nein

C.4 Datenimport

Für jeden in (2) bestimmten Datentyp der Kategorien Bewegungs-, Stamm- und Referenzdaten sollen folgende Fragestellungen geprüft werden:

C.4.1 Quelle

► Aktualisieren sich diese Daten? Gibt es einen Prozess zum regelmäßigen Bezug der Daten?

- Nein, für keine der in (2) genannten Kategorien

► Woher kommen die Daten? - Werden die Daten durch Dritte geliefert? Wenn ja, wer sind die Lieferanten? - Werden die Daten extrahiert? Wenn ja, woher?

- Daten stammen einerseits aus den Bundesländern (BDF-Betreibende) und andererseits aus F&E Projekten
- ▶ Wie zuverlässig ist dieser Prozess? - Gibt es eine (ggf. gesetzliche) Grundlage, die die Zurverfügungstellung oder Lieferung der Daten garantiert? - Sind Änderungen an dieser Grundlage geplant? - Hat das UBA die Möglichkeit Änderungen an der Grundlage vorzunehmen oder darauf einzuwirken?
 - Lieferungen sind nicht zuverlässig, Gesetzesänderungen unter UBA Beteiligung zur Verbesserung sind geplant
- ▶ Gibt es ggf. verschiedene Quellen? (Bei Abweichung zwischen verschiedenen Quellen, sollten die folgenden Punkte je Quelle diskutiert werden.)
 - Ja, s.o.: Bundesländer und F&E

C.4.2 Übermittlungsmodus

Wie werden die Datensätze derzeit übermittelt?

- ▶ Werden die Daten auf Anforderung geliefert bzw. geholt, oder gibt es einen Turnus?
 - Auf Anforderung
- ▶ Wie werden die Daten übermittelt? (etwa per E-Mail, per Upload, per Download)
 - E-Mail, teilweise Download
- ▶ In welchem Format werden die Daten übermittelt?
 - Excel und Access, beide in wechselnden Strukturen
- ▶ Welchen Umfang haben die übermittelte Datensätze? Ist dieser stets gleich? Wird stets der volle Umfang übermittelt oder kann die Übermittlung in mehreren Teilen ggf. zu verschiedenen Zeitpunkten stattfinden?
 - Unterschiedlicher Umfang und abhängig davon, wie viele Bundesländer liefern bzw. wie das Thema des F&E ist; Lieferungen können zu verschiedenen Zeitpunkten stattfinden.
- ▶ Gibt es geplante oder gewünschte Änderungen des Übermittlungsmodus?
 - S.o.

C.4.3 Import in die Fachdatenbank

- ▶ Wie werden die Daten derzeit in die Fachdatenbank übertragen? - Gibt es eine automatische Importroutine im FIS oder werden die Daten manuell übertragen? Mit welchen Werkzeugen? - Müssen die Daten dabei transformiert werden? - Wird stets der gesamte Datensatz importiert oder können auch nur Teile importiert werden?
 - Mit Ausnahme der Datenlieferungen aus drei bayerischen Ämtern, die automatisiert importiert werden, erfolgt der Import jeweils manuell in mehreren Schritten.

C.4.4 Qualitätssicherung

Werden die Daten geprüft? Wenn ja,

- ▶ wann: vor dem Import in die Fachdatenbank oder zu einem späteren Zeitpunkt in der Fachdatenbank?
- ▶ wie: manuell oder automatisch? Mit welchen Werkzeugen?
- ▶ was: syntaktisch, semantisch, inhaltlich?
 - Daten werden erst bei der Auswertung geprüft

Wie ist das Vorgehen bei der Beanstandung und Korrektur der Daten?

- ▶ Werden Fehler selbst korrigiert? Wenn ja, welche Fehlertypen betrifft das und wie wird dabei vorgegangen? Wird das dokumentiert, ggf. wie?
- ▶ Werden Fehler den Datenoriginatoren gemeldet und von diesen korrigiert?
 - Keine Angabe

Wie ist das Vorgehen beim Import von fehlerhaften Datensätzen in die Fachdatenbank?

- ▶ Werden fehlerhafte Datensätze vollständig importiert, wird nur der fehlerfreie Teil importiert, oder wird der Datensatz komplett zurückgewiesen?
- ▶ Bei vollständigem Import der Daten, werden fehlerhafte Daten gekennzeichnet oder nach dem Import wieder gelöscht?
 - Vollständiger Import

Gibt es gewünschte Änderungen?

- ▶ Gibt es QS-Prüfungen, die sich als unnötig oder unpraktikabel erwiesen haben?
- ▶ Gibt es Bedarf für zusätzliche QS-Prüfungen?
 - Es liegen derzeit keine automatisierten oder spezifizierten QS-Prüfverfahren (die über die technische Verarbeitbarkeit hinausgingen) vor. Der Bedarf wird geprüft.

C.4.5 Datenkorrektur und -aktualisierung

Bewegungsdaten: Wie werden Korrekturen von bereits in die Fachdatenbank importierten Datensätzen importiert?

- ▶ Beinhaltet der Korrekturdatensatz auch die bereits fehlerfreien Daten des Originaldatensatzes oder enthält er nur die Korrekturen der fehlerhaften Daten?
- ▶ Werden die Daten überschrieben? Wenn ja, wie werden diese identifiziert?
- ▶ Werden die zuvor importierten Daten vor dem Korrekturimport gelöscht?
 - Die Mechanismen zur Datenkorrektur sehen vor, dass Datenbestände ggf. neu importiert werden und die dadurch ersetzten Datenbestände gelöscht werden. Eine manuelle Korrektur einzelner Werte findet nicht statt.

Stamm- und Referenzdaten: Werden Änderungen bestehender Datensätze berichtet? Wie wird dort beim Import vorgegangen?

- ▶ Werden Stilllegungen erfasst?
- ▶ Werden Vorgänger-Nachfolger-Relationen erfasst?
- ▶ Werden zeitliche Änderung von Eigenschaften erfasst?
 - Es können derzeit nur bayerische Stammdaten automatisiert importiert werden. Das Einrichtungsdatum der Untersuchungsflächen wird verarbeitet. Eine Stilllegung ist bislang nicht vorgesehen.
- ▶ Besteht Bedarf, hierbei von der bisherigen Vorgehensweise abzuweichen?
 - Keine Angabe

C.5 Manuelle Datenänderungen

Können Daten in der Fachdatenbank manuell (also nicht im Rahmen des Datenimports) erzeugt, geändert oder gelöscht werden? Wenn ja:

- ▶ Welche Daten betrifft das?
- ▶ Wird diese Funktionalität genutzt?
- ▶ Woraus resultiert dieser Änderungsbedarf?

Werden die manuellen Änderungen dokumentiert? Können diese rückgängig gemacht werden?

- Eine manuelle Änderung kann theoretisch vorgenommen werden und wird dann in der Historie dokumentiert. Derzeit wird dies nicht genutzt.

Besteht Bedarf, hierbei von der bisherigen Vorgehensweise abzuweichen, neue Möglichkeiten zu schaffen oder ist die bestehende Funktionalität teilweise überflüssig?

- Die bestehende Funktionalität ist für eBIS passend und soll beibehalten werden.

C.6 Datenanalyse

C.6.1 Methodik

Welche Auswertungen werden regelmäßig durchgeführt? Gibt es häufig spezielle, einmalige Auswertungen oder ändert sich der Auswertungsbedarf häufig?

- Beides

Wie werden regelmäßige und unregelmäßige Auswertungen vorgenommen? Manuell oder durch das FIS unterstützt? Welche Werkzeuge kommen zum Einsatz?

- SQL Abfrage nach Bedarf auf FIS

Werden Daten regelmäßig graphisch visualisiert? Wenn ja, wer macht dies und welche Werkzeuge kommen zum Einsatz?

- Fachbetreuer mit R, QGIS, ArcGIS

C.6.2 Berechnungen

Welche aufwändigen Berechnungen sind regelmäßig nötig?

- ▶ zeitliche Aggregationen?
- ▶ geographische Aggregationen?
- ▶ Aggregationen über Parameter? (z.B. Ionenbilanzen, Nitrat und Nitrat-N, o. ä.)?
- ▶ Klassierung von Messwerten? Häufigkeitsverteilungen?
 - alle

Wann finden diese Berechnungen derzeit statt?

- ▶ Bei der Auswertung
 - Bei der Auswertung, Ergebnisse werden nicht in der DB gespeichert. Auswertungen sollen nicht in der Datenbank gespeichert werden
- ▶ Bereits beim Datenimport (Ergebnisse werden in der Datenbank gespeichert)
- ▶ Manuell zu jedem beliebigen Zeitpunkt (Ergebnisse werden in der Datenbank gespeichert)

C.7 Datenexport

Gibt es Bedarf für manuellen Export von Rohdaten zur Weiterverarbeitung außerhalb des FIS?
Wenn ja,

- ▶ In welchem Format? Datenbankdump (SQL), Excel, CSV?
 - Alle, hauptsächlich Excel
- ▶ Welche Datenarten werden exportiert?
 - Auswertungsabhängig
- ▶ Was sind typische Anwendungsfälle?
 - Veränderung der Konzentration eines Elements über die Zeit mit den erklärenden Faktoren
- ▶ Gibt es weitere Anforderungen (etwa Filter)?
 - Ja, auswertungsbedingte Filter nach verschiedenen Kriterien

C.8 Berichterstattung

Welche Berichtsprozesse gibt es derzeit oder wird es künftig geben?

- Im Moment nur INSPIRE, weitere Berichtspflichten könnten sich durch EU und ausgestalteten Koalitionsvertrag ergeben.

Für jeden Berichtsprozess stellen sich folgende Fragen:

- ▶ Wer ist Empfänger des Berichts?
- ▶ In welchem Turnus wird berichtet? Oder wird auf Anforderung berichtet?
- ▶ Auf welcher gesetzlichen Grundlage wird berichtet?
- ▶ In welchem Format wird bisher berichtet? In welchem Format soll/darf berichtet werden?
- ▶ Wie werden die Daten bisher übermittelt? Wie sollen/dürfen die Daten übermittelt werden?
- ▶ Sind für die Erstellung des Berichts externe Daten oder Informationen nötig, die nicht Teil des FIS sind?
- ▶ Wie wird bei Nachforderungen und Korrekturen vorgegangen?

C.9 Verschiedenes

- ▶ Welche Funktionen des FIS wurden in den Vergangenheit nicht oder extrem selten gebraucht?
 - Die Möglichkeit zur Änderung von Messwerten wurden nicht gebraucht, die Möglichkeiten zur Pflege von Untersuchungsflächen selten.
- ▶ Welche im FIS verwaltete Daten wurden in den Vergangenheit nicht oder extrem selten gebraucht?
 - Die Informationen zur Substratsystematik des Profils sowie die Bewirtschaftungs- und Klimainformationen.
- ▶ Welche Funktionen oder Daten fehlten in der Vergangenheit?
 - Automatische Aktualisierung einiger Listen (Klimadaten, Verwaltungsdaten); Für den Import von Daten aus Länderlieferungen bestand bislang keine Möglichkeit, unbekannte Listeneinträge (eBIS umfasst 196 Listen) — soweit möglich — automatisiert zu bekannten Listeneinträgen zu konvertieren.

D Prozessanalyse Grundwasser

D.1 Status quo: Fachdatenbanken und Software

Was für eine Datenbank nutzt das derzeitige FIS?

- ▶ Gibt es mehrere Datenbanken?
- ▶ Gibt es Informationen, die außerhalb von Datenbanken gehalten werden?
- ▶ Wann wurden die Datenmodelle erarbeitet? Wurden diese regelmäßig gepflegt?

Welche Software wird zur Interaktion mit der Fachdatenbank eingesetzt?

- ▶ Gibt es Spezialsoftware? Welche Technologien werden dabei eingesetzt? Wann wurde diese erstellt? Wurde diese regelmäßig gepflegt?
- ▶ Welche Standardsoftware wird regelmäßig eingesetzt?
- ▶ Wird ggf. verschiedene Software für folgende Aufgabenbereiche genutzt?
 - Datenimport
 - Datenexport
 - Auswertungen
 - Berichterstellung

Das im UBA FG II 2.1 (ab hier: FG) genutzte FIS Grundwasser basiert auf einer Postgres-Datenbank, die sämtliche Fachdaten enthält, einer Webanwendung, die einen rudimentären Zugriff auf die Datenbanktabellen, ein Abfragewerkzeug und ein Berichterstellungsmodul umfasst, sowie einer Reihe von Kommandozeilenwerkzeuge zum Import verschiedener Daten.

D.2 Datenarten

Welche Arten von Daten werden in der/den bestehende(n) Fachdatenbank(en) verarbeitet? Gibt es Datenarten, die derzeit noch nicht verarbeitet werden, aber für das neue FIS Wasser-Boden für dieses Umweltmedium berücksichtigt werden sollen?

- ▶ Bewegungsdaten: Messwerte
 - Messwerte werden nach Messstelle und Parameter erfasst.
 - Der Zeitpunkt der Messung wird tagesgenau erfasst.
 - Es wird erfasst, ob die Messung aus einer filtrierten oder einer unfiltrierten Probe genommen werden. Andere Probenahmematrizen gibt es nicht.
 - Die Bestimmungsgrenze des Messverfahrens wird optional erfasst.
 - Es wird entweder ein Messwert oder die Information, dass der Messwert unter der Bestimmungsgrenze liegt, erfasst. In letzterem Fall wird immer auch die Bestimmungsgrenze erfasst.

- Alle Proben sind Einzelproben. Mischproben werden nicht erfasst.

► Stammdaten:

- Messstellen
 - Messstellen sind die geographischen Entitäten, zu denen Messwerte erhoben werden
- Messnetze
 - Messstellen werden zu verschiedenen administrativen Zwecken zu Messnetzen gruppiert.
 - Messstellen können Teil mehrerer oder keines Messnetzes sein.
- Grundwasserkörper
 - Grundwasserkörper sind geographische Flächen, die mit einer Höhenangabe unter der Geländeoberkante ("Horizont") auf einer diskreten, nicht-metrischen Skala (meist 1, 2 oder 3) versehen sind.
 - Zu jedem Punkt auf der Geländeoberkante Deutschlands existiert ein Grundwasserkörper, der senkrecht unter diesem Punkt liegt

► Referenzdaten

- Parameter
 - Quantifizierbare Charakteristiken der Grundwässer, deren Messungen im FIS gespeichert werden.
- Parametersätze
 - Parameter können zu verschiedenen Zwecken zu Parametersätzen gruppiert werden.
- Physikalische Größen
 - physikalische Größen sind die möglichen Kategorien, in denen Parameter quantifiziert werden können.
- Einheiten
 - physikalische Einheit, mit der eine bestimmte physikalische Größe quantifiziert werden kann.
- Plausibilitäts- und Gütegrenzen
 - Ober- oder Untergrenzen für Messwerte bestimmter Parameter, die entweder bei der Import-QS oder nachgelagerter QS zur Prüfung der Plausibilität von Messwerten oder bei Auswertungen zur Prüfung der Güte des Grundwassers genutzt werden

► Metadaten

- Im FIS Grundwasser werden keine Metadaten zum Datenimport oder zur Datenquelle gespeichert.

- Im neu zu entwickelnden FIS soll analog zu der Vorgehensweise in den anderen Umweltkompartimenten jeder Importvorgang (mit Datum, Quelle und QS-Ergebnissen) protokolliert werden und von jedem Messwert eine eindeutige Zuordnung zum Importvorgang gespeichert werden. Damit soll u.a. die Rückgängigmachung eines beliebigen Importvorgang ermöglicht werden.

D.3 Datenmenge

- ▶ Wie umfangreich sind die Datenbestände?
- ▶ Um welchen Umfang wachsen die Datenbestände regelmäßig?
- ▶ Werden wesentliche Änderungen des Umfangs der Datenbestände erwartet?
 - Stammdaten: In der FIS-Datenbank befinden sich derzeit:
 - 1263 Messstellen
 - 4 Messnetze
 - 1178 Grundwasserkörper
 - Die Anzahl der Messstellen wächst kontinuierlich, es wird aber kein Zuwachs in einer anderen Größenordnung erwartet.
 - Referenzdaten:
 - 1694 Parameter
 - keine Parametersätze
 - 14 physikalische Größen
 - 22 Einheiten
 - keine Plausibilitäts- oder Gütegrenzen
 - Es werden keine wesentlichen Änderungen der Größenordnung des Umfangs der Datenbestände erwartet.
 - Bewegungsdaten:
 - ca. 0,88 Millionen Messwerte
 - Die Menge der Messwerte wächst naturgemäß kontinuierlich. Die Anzahl der Messwerte pro Jahr und Messstelle beträgt über die Jahre 2016 bis 2020 im Mittel 385. Es wird davon ausgegangen, dass auch diese Zahl steigen kann, sich aber in einer vergleichbaren Größenordnung bewegt.

D.4 Datenimport

D.4.1 Datenimport Messstellen

Quelle

Die Informationsquelle über neue Messstellen und Stilllegungen alter Messstellen sind die Umweltbehörden der Bundesländer.

Übermittlungsmodus

Im Rahmen der jährlichen Messdatenübermittlung, ausnahmsweise auch außerhalb dieses Rahmens, werden durch die Umweltbehörden der Bundesländer bei Bedarf neue Messstellen und Stilllegungen alter Messstellen als Excel-Datei per E-Mail an das FG übermittelt. Die Struktur basiert auf einer etablierten Vorlage, weicht aber hinsichtlich des Umfangs der übermittelten Daten regelmäßig von dieser ab.

Weil Messstellenübermittlungen nur vereinzelt stattfinden und die Heterogenität der Datenstrukturen groß ist, wird von einer Automatisierung des Übermittlungsprozesses abgesehen.

Import in die Fachdatenbank

Neue Messstellen werden manuell über die Weboberfläche des FIS eingetragen. Datensätze können teilweise angelegt werden und nachträglich bearbeitet werden. Stilllegungen werden derzeit nicht vermerkt, sollen aber künftig vermerkt werden.

Da der Übermittlungsprozesses nicht verändert werden soll, soll auch der Import weiterhin über eine Weboberfläche des FIS erfolgen.

Qualitätssicherung

Bei der Eintragung und bei der Änderung von Messstellen in der Weboberfläche des FIS, wird derzeit folgendes automatisch geprüft:

- ▶ Eindeutigkeit der Messstellennummer für das Bundesland
- ▶ syntaktische Korrektheit der WFD- und EIONET-Messstellennummern gemäß des durch die EUA vorgegebenen Schemas,
- ▶ Richtigkeit der angegebenen Koordinate in Bezug auf das angegebene Bundesland und den angegebenen Grundwasserkörper,
- ▶ Filteroberkante > Filterunterkante

Datensätze, die diese Prüfung nicht bestehen, können nicht angelegt werden. Existierende Datensätze können nicht so verändert werden, dass dies zum Nichtbestehen dieser Prüfungen führt.

Dieses Vorgehen sowie der Umfang der QS-Prüfungen sollen beibehalten werden.

Datenkorrektur und -aktualisierung

Korrekturen werden auf demselben Weg wie initiale Daten übermittelt. Der Korrekturdatensatz enthält die kompletten Messstelleninformationen, nicht nur die fehlerhaften Daten. Die Daten werden manuell in der Weboberfläche editiert.

Messstellen werden bei Stilllegung nicht aus dem Datenbestand gelöscht, sondern sollen als stillgelegt markiert werden. Bei stillgelegten Messstellen soll der Grund der Stilllegung und wenn vorhanden, die Ersatzmessstelle erfasst werden.

Die Veränderung von Attributen der Messstellen wird nicht chronologisch erfasst.

D.4.2 Datenimport Messnetze

Messnetze sind administrative Gruppierungen von Messstellen zu Auswertungszwecken und werden ausschließlich vom FG über die Weboberfläche des FIS angelegt und verwaltet.

D.4.3 Datenimport Grundwasserkörper

Quelle

Die vom FG verwendete Liste der Grundwasserkörper wird durch die Bundesanstalt für Gewässerkunde erstellt und alle sechs Jahre aktualisiert.

Übermittlungsmodus

Die Grundwasserkörperliste kann im Portal "Wasserblick" der BfG als Shape-Datei heruntergeladen werden.

Die Liste enthält stets alle gültigen Grundwasserkörper. Eine eindeutige Zuordnung von Grundwasserkörpern zu Grundwasserkörpern der Vorversion einer Liste ist nicht möglich.

Import in die Fachdatenbank

Die Liste wird derzeit automatisiert mit einem hierfür entwickelten Kommandozeilenwerkzeug transformiert und importiert. Einen Mechanismus zur Erkennung und Zuordnung der Grundwasserkörper bei Import einer neuen Version der Liste gibt es nicht.

Der Importprozess soll auch zukünftig mit einem geeigneten Kommandozeilenwerkzeug vorgenommen werden, das weiterhin befähigt werden soll, Grundwasserkörper zum großen Teil automatisch zu erkennen und zuzuordnen.

Qualitätssicherung

Beim Import wird derzeit folgendes automatisch geprüft:

- ▶ syntaktische Korrektheit der WFD-Nummer des Wasserkörper gemäß des durch die EUA vorgegebenen Schemas,
- ▶ Horizontnummer muss eine positive Ganzzahl sein.

Dies soll beibehalten werden.

Datenkorrektur und -aktualisierung

Einen Korrekturprozess gibt es nicht. Für die Aktualisierung soll ein Prozess entwickelt werden, der Stilllegung und Vorgänger-Nachfolger-Relationen erkennen und speichern kann.

D.4.4 Datenimport Parameter

Quelle

Die vom FG verwendete Liste der Parameter wird durch das Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV) gepflegt. Darin werden Parameter mit Nummern kodiert.

Im FIS wird jeder Parameter zusätzlich mit der physikalischen Größe, die er beschreibt, und der Standardeinheit zur Darstellung von Werten versehen.

Übermittlungsmodus

Die Liste wird dem FG manuell im Excel-Format übermittelt.

Die Liste enthält stets alle gültigen Parameternummern. Eine eindeutige Zuordnung von Parametern zu Parametern der Vorversion einer Liste ist über die Parameternummer möglich. Die Liste wird regelmäßig erweitert, bestehende Einträge werden jedoch nicht verändert oder gelöscht.

Import in die Fachdatenbank

Die Liste wird derzeit automatisiert mit einem hierfür entwickelten Kommandozeilenwerkzeug transformiert und importiert.

Der Importprozess soll über einen Dateiupload künftig über die Weboberfläche durchgeführt werden können. Dabei sollen nur die in der Liste neu vorkommenden Parameter erkannt und importiert werden.

Physikalische Größen und Standardeinheiten von Parametern sollen über die Weboberfläche bearbeitet werden, sofern noch keine Messwerte zu diesen Parametern existieren.

Qualitätssicherung

Die Anwendung prüft, dass die Parameternummer eindeutig ist, und dass die gewählte Standardeinheit für die gewählte physikalische Größe gültig ist.

Datenkorrektur und -aktualisierung

Eine manuelle Korrektur und Bearbeitung von Parametern soll ermöglicht werden, sofern noch keine Messwerte zu diesen Parametern existieren. Andernfalls ist lediglich der Parametername änderbar.

Entfernungen oder Vorgänger-Nachfolger-Relationen von Parametern sind nicht vorgesehen.

D.4.5 Datenimport Parametergruppen

Parametergruppen sind administrative Gruppierungen von Parametern zu Auswertungszwecken und werden ausschließlich vom FG über die Weboberfläche des FIS angelegt und verwaltet.

D.4.6 Datenimport physikalische Größen & Einheiten

Physikalische Größen und Einheiten sollen ausschließlich über die Weboberfläche konfiguriert werden. Dabei dürfen existierende Größen und Einheiten nicht gelöscht werden, wenn sie in Verwendung sind.

D.4.7 Datenimport Plausibilität- und Gütegrenzen

Plausibilität- und Gütegrenzen sollen ausschließlich über die Weboberfläche konfiguriert erstellt, bearbeitet und gelöscht werden.

D.4.8 Datenimport Messwerte

Quelle

Messwerte werden von den 16 Landesumweltbehörden übermittelt.

Die rechtliche Grundlage für die Übermittlung der Messwerte ist eine Verwaltungsvereinbarung zwischen Bund und Ländern. Änderungen an der Grundlage sind nicht geplant.

Übermittlungsmodus

Die Übermittlung erfolgt per E-Mail.

Die Übermittlung aller Messwerte eines Kalenderjahres erfolgt auf Anforderung mindestens im jährlichen Turnus.

In der Regel werden die Messwerte in einer, manchmal in mehreren Dateien übermittelt. Alle Lieferungen folgen einer durch die Verwaltungsvereinbarung vorgegebene Tabellenstruktur, die in Festbreiten-Text-, CSV- oder Exceldateien abgebildet wird. Die Tabellenstruktur wird nicht immer eingehalten, insbesondere der Spaltenumfang, die Spaltenkopfbeschreibungen und die Spaltenreihenfolge kann abweichen. Auch innerhalb einer Lieferung mit mehreren Dateien stimmt das Format nicht immer zwischen allen Dateien überein.

Für die Messwertübermittlung soll ein standardisiertes XML-Format entwickelt werden. Die bestehenden Formate sollen aber mittelfristig weiter unterstützt werden.

Import in die Fachdatenbank

Die Messwerte werden derzeit automatisiert mit je nach Bundesland unterschiedlichen, hierfür entwickelten Kommandozeilenwerkzeugen transformiert und importiert.

Der Import soll künftig durch die Lieferanten über die Weboberfläche des externen Anwendungsteils vorgenommen werden.

Beim Import wird der Messwert automatisch in die dem Parameter im FIS zugeordneter Standardeinheit umgerechnet.

Qualitätssicherung

Die Daten werden derzeit beim automatisierten Import wie folgt geprüft:

- ▶ Pflichtfelder müssen vorhanden sein, Datentyp muss gültig sein.
- ▶ Das Datum muss gültig sein und in dem korrekten Berichtszeitraum liegen.
- ▶ Die Parameternummer muss gültig sein.
- ▶ Die Einheitennummer muss gültig sein.
- ▶ Die Messstellenummer muss für das Bundesland gültig sein.
- ▶ Messwert und Bestimmungsgrenze müssen numerisch sein.
- ▶ Die Bestimmungsgrenze muss angegeben werden, wenn der Messwert unter der Bestimmungsgrenze liegt.
- ▶ Der Messwert darf nicht angegeben werden, wenn der Wert unter der Bestimmungsgrenze liegt.
- ▶ Der Messwert muss angegeben werden, solange er nicht unter der Bestimmungsgrenze liegt.

- ▶ Der Messwert darf nicht unter der Bestimmungsgrenze liegen.
- ▶ Die Trennkennzahl muss ein 1 oder 6 sein für alle Parameter, die in der Probe gemessen werden. Andernfalls darf keine Trennkennzahl angegeben werden.
- ▶ Die angegebene Einheit muss zu der selben physikalischen Größe wie der angegebene Parameter gehören.
- ▶ Pro Messstelle, Parameter, Datum und Trennkennzahl darf nur ein Messwert existieren

Durch die in den Kommandozeilenwerkzeugen bestehenden Transformationsprozesse werden einige Fehler derzeit automatisch korrigiert. Dies soll künftig nicht stattfinden. Alle Fehler sollen stattdessen für die Lieferanten zur Korrektur aufbereitet werden.

Eine Datei wird nur importiert, wenn sie vollständig frei von schweren Fehlern (Blocker) ist. Dies soll auch künftig so gehandhabt werden.

Es sollen künftig QS-Routinen ausgearbeitet werden, die die Plausibilität der Messwerte auf Grundlage explizit definierter Grenzen oder im Vergleich mit vergangenen Messwerten der gleichen Messstelle prüft. Diese Routinen sollen aber nur Warnungen ausgeben, die den Import nicht verhindern.

Enthält eine Datenlieferung für eine Messstelle aus dem Bundesland des Lieferanten keine Messwerte und ist die Messstelle nicht als stillgelegt markiert, soll dies auch als Warnung gemeldet werden.

Enthält eine Datenlieferung für eine Messstelle für bestimmte als besonders wichtig erachtete Parameter keine Daten, soll dies auch als Warnung gemeldet werden.

Die Plausibilität der Messwerte von Ionenkonzentrationen soll durch eine Ionenbilanzierung und Abgleich mit der elektr. Leitfähigkeit geprüft werden und im Fall größerer Abweichung bei der Bilanz als Warnung gemeldet werden.

Datenkorrektur und -aktualisierung

Der Korrekturdatensatz soll stets auch die bereits fehlerfreien Daten des Originaldatensatzes enthalten. Beim Import eines Korrekturdatensatzes soll der Lieferant den früheren Liefervorgang auswählen, für den die neue Lieferungen eine Korrektur darstellt. Im Falle eines erfolgreichen Imports des Korrekturdatensatzes sollen dabei automatisch alle Messwerte, die durch die frühere Lieferung importiert wurden, gelöscht werden.

D.5 Manuelle Datenänderungen

Messstellen und Messnetze sollen über die Weboberfläche erstellt und bearbeitet werden können. Die Änderungen sollen dabei protokolliert werden. Grundwasserkörper sollen nicht außerhalb der Importprozesses bearbeitet werden.

Parameter, physikalische Größen und Einheiten sollen über die Weboberfläche erstellt und bearbeitet werden.

Es gibt keinen Bedarf für die manuelle Änderung von Messwerten. Korrekturen sollen ausschließlich durch die Korrekturfunktion der Datenlieferung im externen Anwendungsteil vorgenommen werden.

D.6 Datenanalyse

D.6.1 Methodik

Komplexe Analysen werden durch das FG hauptsächlich mit Statistiksoftware wie R vorgenommen. Um das zu ermöglichen, soll eine Funktion zur automatischen Generierung eines geeigneten SQL-Dumps erstellt werden. (siehe unten, Datenexport)

Zur einfachen Abfrage von Daten soll eine Abfragewerkzeug in der Weboberfläche des FIS die Abfrage von Messwerten mit zeitlicher und räumlicher Einschränkung sowie Einschränkung nach Parametern, Messstellen oder deren Eigenschaften möglich sein.

Die Ergebnisse sollen mit folgenden Visualisierungsmethoden aufbereitet werden:

- ▶ Kurvendiagramm für zeitliche Verläufe von Messwerten mit eine Kurve pro Parameter an einer Messstelle oder über mehrere Messstellen aggregiert
- ▶ Kurvendiagramm für zeitliche Verläufe von Messwerten mit eine Kurve pro Messstelle oder Grundwasserkörper für einen Parameter
- ▶ Histogramme, Boxplots oder andere Visualisierungsmethoden der deskriptiven Statistik für die Verteilung von Messwerten eines Parameters an einer Messstelle oder über mehrere Messstellen aggregiert, inklusive Darstellung von Extrema, Perzentilen, Interquartilsabstand, z-Score etc.
- ▶ Zeitstrahl der Messzeitpunkte an einer oder mehreren Messstellen

Bei Auswertungen soll sowohl grundsätzlich als im Einzelfall abweichend entschieden werden können, wie Messwerte, die unterhalb einer Bestimmungsgrenze liegen, verrechnet werden (Optionen sind als Nullwert, als halber Bestimmungsgrenzwert, diese Werte ausschließen).

Nach Möglichkeit sollen einzelne Diagramme auch über die Schnellauswahl einer Messstelle außerhalb des Abfragewerkzeug gezeigt werden.

D.6.2 Berechnungen

Sämtliche in den Auswertungsroutinen notwendige Aggregationen sollen automatisch bei der Auswertung vorgenommen werden.

D.7 Datenexport

Messwertdaten sollen zum Zweck der weiteren Analyse mit externer Analysesoftware exportierbar sein.

Die zu exportierende Datenmenge soll dabei mit verschiedenen Filtermethoden zeitlich, nach Parameter(n) und nach Messstellen und verschiedenen Eigenschaften eingeschränkt werden können.

Die Daten sollen im Excel, CSV, JSON und einem SQL-Datenbankdump-Format exportiert werden. Der Export soll über die Weboberfläche oder mittels einem programmatischen Zugriff über eine HTTP-API erfolgen können.

D.8 Berichterstattung

Für das WISE-6 water quality reporting an die Europäische Umweltagentur (EUA) müssen in jährlichem Turnus Messwerte des vorvergangenen Kalenderjahres berichtet werden.

Der Bericht wird dazu in einem definierten Excel- oder XML-Format auf einer EUA-Webseite hochgeladen.

Im aktuellen FIS Grundwasser ist dazu eine Funktion zur Erzeugung und zum Download einer geeigneter XML-Datei implementiert, die weitergenutzt werden soll.

Für die Berichterstellung sind folgende Informationen nötig, die über die Weboberfläche konfiguriert werden sollen:

- ▶ Umfang der zu berichtenden Parameter
- ▶ Für jeden zu berichtenden Parameter:
 - Code des Parameters, der von WISE-6 zur Identifikation verwendet wird.
 - Einheit, in der Messwerte dieses Parameters berichtet werden sollen
 - Bezeichnung dieser Einheit, die von WISE-6 zur eindeutigen Identifikation der Einheit verwendet wird.

Beim Upload zur EUA-Webseite findet eine Reihe von dokumentierten QS-Prüfungen statt, die im FIS bei der Berichterstellung ebenso geprüft werden sollen, sofern sie nicht schon Teil der Import-QS sind.

E Prozessanalyse Seen

E.1 Status quo: Fachdatenbanken und Software

Das im UBA FG II 2.4 (ab hier: FG) genutzte FIS Seen basiert auf einer Postgres-Datenbank, die sämtliche Fachdaten enthält und einer Webanwendung, die u.a. einen schreibenden Zugriff auf Messwerte und Stammdaten erlaubt, ein Abfragewerkzeug, ein Import und ein Berichterstellungsmodul umfasst.

E.2 Datenarten

E.2.1 Stammdaten

► Messstellen

- Messstellen sind die geographischen Entitäten, zu denen Messwerte erhoben werden. Messstellen liegen in Seen und sind diesen eindeutig zuzuordnen.

► Messnetze

- Messstellen werden zu verschiedenen administrativen Zwecken zu Messnetzen gruppiert.
- Messstellen können Teil mehrerer oder keines Messnetzes sein.

► Seen

- Seen werden mit ihrer geographischen Ausdehnung (Shape), verschiedenen geographischen und hydrologischen Eigenschaften, sowie geographischen und statistischen Informationen über ihr Einzugsgebiet und dessen Flächennutzung erfasst.
- Seen sind Flussgebietseinheiten eindeutig zuzuordnen.

► Flussgebietseinheiten

- Flussgebietseinheiten bilden eine lückenlose und überlappungsfreie Überdeckung des Staatsgebiets Deutschlands.

E.2.2 Referenzdaten

► Parameter

- Quantifizierbare Charakteristiken der Seewässer, deren Messungen im FIS gespeichert werden.

► Parametersätze

- Parameter können zu verschiedenen Zwecken zu Parametersätzen gruppiert werden.

► Physikalische Größen

- physikalische Größen sind die möglichen Kategorien, in denen Parameter quantifiziert werden können.
- ▶ Einheiten
 - physikalische Einheit, mit der eine bestimmte physikalische Größe quantifiziert werden kann.
- ▶ Plausibilitäts- und Gütegrenzen
 - Ober- oder Untergrenzen für Messwerte bestimmter Parameter, die entweder bei der Import-QS oder nachgelagerter QS zur Prüfung der Plausibilität von Messwerten oder bei Auswertungen zur Prüfung der Güte des Grundwassers genutzt werden.

E.2.3 Bewegungsdaten

- ▶ Messwerte:
 - Messwerte werden nach Messstelle und Parameter erfasst.
 - Der Zeitpunkt der Messung wird tagesgenau erfasst.
 - Die Matrix der Probe wird erfasst. Die Probe kann entweder aus filtriertem oder unfiltriertem Wasser, einer Muschel oder einem Fischmuskel genommen werden.
 - Messwerte können von Einzel- oder von Mischproben aus Proben aus verschiedenen Tiefen bestimmt werden. (Mischproben aus Proben, die an unterschiedlichen Tagen genommen wurden, werden nicht erfasst.)
 - Die Tiefe der Probenahme wird erfasst. Im Fall von Mischproben wird die untere und die obere Grenze der Tiefen erfasst, zwischen denen die Einzelproben genommen wurden, erfasst.
 - Die Bestimmungsgrenze des Messverfahrens wird optional erfasst.
 - Es wird entweder ein Messwert oder die Information, dass der Messwert unter der Bestimmungsgrenze liegt, erfasst. In letzterem Fall wird immer auch die Bestimmungsgrenze erfasst.

E.2.4 Metadaten

Im FIS Seen werden keine Metadaten zum Datenimport oder zur Datenquelle gespeichert.

Im neu zu entwickelnden FIS soll analog zu der Vorgehensweise in den anderen Umweltkompartimenten jeder Importvorgang (mit Datum, Quelle und QS-Ergebnissen) protokolliert werden und von jedem Messwert eine eindeutige Zuordnung zum Importvorgang gespeichert werden. Damit soll u.a. die Rückgängigmachung eines beliebigen Importvorgang ermöglicht werden.

E.3 Datenmenge

E.3.1 Stammdaten

In der FIS-Datenbank befinden sich derzeit:

- ▶ 81 Messstellen
- ▶ 4 Messnetze
- ▶ 73 Seen
- ▶ 10 Flussgebietseinheiten

Die Anzahl der Messstellen und Seen kann steigen, es wird aber kein Zuwachs in einer anderen Größenordnung erwartet.

E.3.2 Referenzdaten

- ▶ 1602 Parameter
- ▶ 15 Parametersätze (disjunkte “Obergruppen”)
- ▶ 22 Einheiten
- ▶ 64 Gütegrenzen (“Umweltqualitätsnorm (UQN)”)

Es werden keine wesentlichen Änderungen der Größenordnung des Umfangs der Datenbestände erwartet.

E.3.3 Bewegungsdaten

- ▶ ca. 0,26 Millionen Messwerte

Die Menge der Messwerte wächst naturgemäß kontinuierlich. Die Anzahl der Messwerte pro Jahr und Messstelle beträgt über die Jahre 2007 bis 2011 im Mittel 592. Es wird davon ausgegangen, dass auch diese Zahl steigen kann, sich aber in einer vergleichbaren Größenordnung bewegt.

E.4 Datenimport

E.4.1 Datenimport Messstellen

Quelle

Die Informationsquelle über neue Messstellen und Stilllegungen alter Messstellen sind die Umweltbehörden der Bundesländer.

Übermittlungsmodus

Im Rahmen der jährlichen Messdatenübermittlung, ausnahmsweise auch außerhalb dieses Rahmens, werden durch die Umweltbehörden der Bundesländer bei Bedarf neue Messstellen und

Stilllegungen alter Messstellen als Excel-Datei per E-Mail an das FG übermittelt. Die Struktur basiert auf einer etablierten Vorlage, weicht aber hinsichtlich des Umfangs der übermittelten Daten regelmäßig von dieser ab.

Weil Messstellenübermittlungen nur vereinzelt stattfinden und die Heterogenität der Datenstrukturen groß ist, wird von einer Automatisierung des Übermittlungsprozesses abgesehen.

Import in die Fachdatenbank

Neue Messstellen werden manuell über die Weboberfläche des FIS eingetragen. Datensätze können teilweise angelegt werden und nachträglich bearbeitet werden. Stilllegungen werden derzeit nicht vermerkt, sollen aber künftig vermerkt werden.

Da der Übermittlungsprozesses nicht verändert werden soll, soll auch der Import weiterhin über eine Weboberfläche des FIS erfolgen.

Qualitätssicherung

Bei der Eintragung und bei der Änderung von Messstellen in der Weboberfläche des FIS, wird derzeit folgendes automatisch geprüft:

- ▶ Eindeutigkeit der Messstellenummer für das Bundesland
- ▶ Korrektheit der Datentypen, z.B. Gültigkeit von Datumsangaben, Koordinaten, Fließkommazahlen

Weiterhin soll folgendes künftig geprüft werden:

- ▶ Richtigkeit der angegebenen Koordinate in Bezug auf das angegebene Bundesland und den angegebenen See
- ▶ syntaktische Korrektheit der WFD-Messstellenummern gemäß des durch die EUA vorgegebenen Schemas

Datensätze, die diese Prüfung nicht bestehen, können nicht angelegt werden. Existierende Datensätze können nicht so verändert werden, dass dies zum Nichtbestehen dieser Prüfungen führt. Dieses Vorgehen soll beibehalten werden.

Datenkorrektur und -aktualisierung

Korrekturen werden auf demselben Weg wie initiale Daten übermittelt. Der Korrekturdatensatz enthält die kompletten Messstelleninformationen, nicht nur die fehlerhaften Daten. Die Daten werden manuell in der Weboberfläche editiert.

Messstellen werden bei Stilllegung nicht aus dem Datenbestand gelöscht, sondern sollen als stillgelegt markiert werden. Bei stillgelegten Messstellen soll der Grund der Stilllegung und wenn vorhanden, die Ersatzmessstelle erfasst werden.

E.4.2 Datenimport Messnetze

Messnetze sind administrative Gruppierungen von Messstellen zu Auswertungszwecken und werden ausschließlich vom FG über die Weboberfläche des FIS angelegt und verwaltet.

E.4.3 Datenimport Seen

Quelle

Seen werden im FIS nur angelegt, wenn es zugehörige Messstellen gibt. Die Informationsquelle über neue anzulegende Seen sind daher ebenfalls die Umweltbehörden der Bundesländer.

Übermittlungsmodus

Der Übermittlungsmodus gleicht dem Modus für die Übermittlung von Messstellen.

Weil auch Seenübermittlungen nur vereinzelt stattfinden, wird von einer Automatisierung des Übermittlungsprozesses abgesehen.

Import in die Fachdatenbank

Neue Seen werden manuell über die Weboberfläche des FIS eingetragen. Datensätze können teilweise angelegt werden und nachträglich bearbeitet werden. Stilllegungen gibt es nicht.

Da der Übermittlungsprozesses nicht verändert werden soll, soll auch der Import weiterhin über eine Weboberfläche des FIS erfolgen.

Qualitätssicherung

Bei der Eintragung und bei der Änderung von Seen in der Weboberfläche des FIS, wird derzeit folgendes automatisch geprüft:

- ▶ Eindeutigkeit des Namens und der Kurzbeschreibung des Sees
- ▶ Korrektheit der Datentypen, z.B. Gültigkeit von Datumsangaben, Koordinaten, Fließkommazahlen

Weiterhin soll folgendes künftig geprüft werden:

- ▶ Richtigkeit der angegebenen Seefläche in Bezug auf die angegebene Flussgebietseinheit
- ▶ logisch-arithmetische Prüfungen der Seeigenschaften (z. B. mittlere Tiefe \leq maximale Tiefe)

Datensätze, die diese Prüfung nicht bestehen, können nicht angelegt werden. Existierende Datensätze können nicht so verändert werden, dass dies zum Nichtbestehen dieser Prüfungen führt. Dieses

Vorgehen soll beibehalten werden.

Datenkorrektur und -aktualisierung

Korrekturen werden auf demselben Weg wie initiale Daten übermittelt. Der Korrekturdatensatz enthält die kompletten Seeinformationen, nicht nur die fehlerhaften Daten. Die Daten werden manuell in der Weboberfläche editiert.

E.4.4 Datenimport Flussgebietseinheiten

Die für das FG maßgebliche Liste der Flussgebietseinheiten wird durch die EU-Wasserrahmenrichtlinie vorgegeben. Die Liste ändert sich nicht, ein Importprozess oder Editier-Prozess wird nicht benötigt.

E.4.5 Datenimport Parameter

Quelle

Die vom FG verwendete Liste der Parameter wird durch das Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV) gepflegt. Darin werden Parameter mit Nummern kodiert.

Im FIS soll jeder Parameter zusätzlich mit der physikalischen Größe, die er beschreibt, und der Standardeinheit zur Darstellung von Werten versehen.

Übermittlungsmodus

Die Liste wird dem FG manuell im Excel-Format übermittelt.

Die Liste enthält stets alle gültigen Parameternummern. Eine eindeutige Zuordnung von Parametern zu Parametern der Vorversion einer Liste ist über die Parameternummer möglich. Die Liste wird regelmäßig erweitert, bestehende Einträge werden jedoch nicht verändert oder gelöscht.

Import in die Fachdatenbank

Die Liste wird derzeit manuell bearbeitet.

Der Importprozess soll über einen Dateiupload künftig über die Weboberfläche durchgeführt werden können. Dabei sollen nur die in der Liste neu vorkommenden Parameter erkannt und importiert werden.

Physikalische Größen und Standardeinheiten von Parametern sollen über die Weboberfläche bearbeitet werden, sofern noch keine Messwerte zu diesen Parametern existieren.

Qualitätssicherung

Die Anwendung soll prüfen, dass die Parameternummer eindeutig ist, und dass die gewählte Standardeinheit für die gewählte physikalische Größe gültig ist.

Datenkorrektur und -aktualisierung

Eine manuelle Korrektur und Bearbeitung von Parametern soll ermöglicht werden, sofern noch keine Messwerte zu diesen Parametern existieren. Andernfalls ist lediglich der Parametername abänderbar.

Entfernungen oder Vorgänger-Nachfolger-Relationen von Parametern sind nicht vorgesehen.

E.4.6 Datenimport Parametersätze

Messnetze sind administrative Gruppierungen von Parametern zu Auswertungszwecken und werden ausschließlich vom FG über die Weboberfläche des FIS angelegt und verwaltet.

E.4.7 Datenimport physikalische Größen & Einheiten

Physikalische Größen und Einheiten sollen ausschließlich über die Weboberfläche konfiguriert werden. Dabei dürfen existierende Größen und Einheiten nicht gelöscht werden, wenn sie in Verwendung sind.

E.4.8 Datenimport Plausibilität- und Gütegrenzen

Plausibilität- und Gütegrenzen sollen ausschließlich über die Weboberfläche konfiguriert erstellt, bearbeitet und gelöscht werden.

E.4.9 Datenimport Messwerte

Quelle

Messwerte werden von den 16 Landesumweltbehörden übermittelt.

Die rechtliche Grundlage für die Übermittlung der Messwerte ist eine Verwaltungsvereinbarung zwischen Bund und Ländern. Änderungen an der Grundlage sind nicht geplant.

Übermittlungsmodus

Die Übermittlung erfolgt per E-Mail.

Die Übermittlung aller Messwerte eines Kalenderjahres erfolgt auf Anforderung mindestens im jährlichen Turnus.

In der Regel werden die Messwerte in einer, manchmal in mehreren Dateien übermittelt. Alle Lieferungen folgen einer durch die Verwaltungsvereinbarung vorgegebene Tabellenstruktur, die in Festbreiten-Text-, CSV- oder Exceldateien abgebildet wird. Die Tabellenstruktur wird nicht immer eingehalten, insbesondere der Spaltenumfang, die Spaltenkopfbeschreibungen und die Spaltenreihenfolge kann abweichen. Auch innerhalb einer Lieferung mit mehreren Dateien stimmt das Format nicht immer zwischen allen Dateien überein.

Für die Messwertübermittlung soll ein standardisiertes XML-Format entwickelt werden. Die bestehenden Formate sollen aber mittelfristig weiter unterstützt werden.

Import in die Fachdatenbank

Die Messwerte werden derzeit automatisiert mit je nach Bundesland unterschiedlichen, hierfür entwickelten Kommandozeilenwerkzeugen transformiert und importiert.

Der Import soll künftig durch die Lieferanten über die Weboberfläche des externen Anwendungsteils vorgenommen werden.

Beim Import wird der Messwert automatisch in die dem Parameter im FIS zugeordneter Standardeinheit umgerechnet.

Qualitätssicherung

Die Daten sollen beim Import automatisiert wie folgt geprüft werden:

- ▶ Pflichtfelder müssen vorhanden sein, Datentyp muss gültig sein.
- ▶ Das Datum muss gültig sein und dem korrekten Berichtszeitraum liegen.
- ▶ Die Parameternummer muss gültig sein.
- ▶ Die Einheitennummer muss gültig sein.
- ▶ Die Messstellenummer muss für das Bundesland gültig sein.
- ▶ Messwert und Bestimmungsgrenze müssen numerisch sein.
- ▶ Die Bestimmungsgrenze muss angegeben werden, wenn der Messwert unter der Bestimmungsgrenze liegt.

- ▶ Der Messwert darf nicht angegeben werden, wenn der Wert unter der Bestimmungsgrenze liegt.
- ▶ Der Messwert muss angegeben werden, solange er nicht unter der Bestimmungsgrenze liegt.
- ▶ Der Messwert darf nicht unter der Bestimmungsgrenze liegen.
- ▶ Die Trennkennzahl muss ein 1 oder 6 sein für alle Parameter, die in der Probe gemessen werden. Andernfalls darf keine Trennkennzahl angegeben werden.
- ▶ Die angegebene Einheit muss zu derselben physikalischen Größe wie der angegebene Parameter gehören.
- ▶ Pro Messstelle, Parameter, Datum und Trennkennzahl darf nur ein Messwert existieren.
- ▶ Die Probetiefegrenzen müssen angegeben und nicht-negativ sein, bei Einzelproben muss die obere Grenze der unteren Grenze entsprechen. Andernfalls darf die obere Grenze nicht größer als die untere Grenze sein.

Eine automatische Fehlerkorrektur soll nicht stattfinden. Alle Fehler sollen stattdessen für die Lieferanten zur Korrektur aufbereitet werden.

Eine Datei soll nur importiert werden, wenn sie vollständig frei von schweren Fehlern (Blocker) ist.

Es sollen künftig QS-Routinen ausgearbeitet werden, die die Plausibilität der Messwerte auf Grundlage explizit definierter Grenzen oder im Vergleich mit vergangenen Messwerten der gleichen Messstelle prüft. Diese Routinen sollen aber nur Warnungen ausgeben, die den Import nicht verhindern.

Enthält eine Datenlieferung für eine Messstelle aus dem Bundesland des Lieferanten keine Messwerte und ist die Messstelle nicht als stillgelegt markiert, soll dies auch als Warnung gemeldet werden.

Enthält eine Datenlieferung für eine Messstelle für bestimmte als besonders wichtig erachtete Parameter keine Daten, soll dies auch als Warnung gemeldet werden.

Datenkorrektur und -aktualisierung

Der Korrekturdatensatz soll stets auch die bereits fehlerfreien Daten des Originaldatensatzes enthalten. Beim Import eines Korrekturdatensatzes soll der Lieferant den früheren Liefervorgang auswählen, für den die neue Lieferungen eine Korrektur darstellt. Im Falle eines erfolgreichen Imports des Korrekturdatensatzes sollen dabei automatisch alle Messwerte, die durch die frühere Lieferung importiert wurden, gelöscht werden.

E.5 Manuelle Datenänderungen

Messstellen, Seen und Messnetze sollen über die Weboberfläche erstellt und bearbeitet werden können. Die Änderungen sollen dabei protokolliert werden. Flussgebietseinheiten sollen nicht bearbeitet werden.

Parameter, physikalische Größen und Einheiten sollen über die Weboberfläche erstellt und bearbeitet werden.

Es gibt keinen Bedarf für die manuelle Änderung von Messwerten. Korrekturen sollen ausschließlich durch die Korrekturfunktion der Datenlieferung im externen Anwendungsteil vorgenommen werden.

E.6 Datenanalyse

E.6.1 6.1 Methodik

Komplexe Analysen werden durch das FG hauptsächlich mit Statistiksoftware wie R vorgenommen. Um das zu ermöglichen, soll eine Funktion zur automatischen Generierung eines geeigneten SQL-Dumps erstellt werden. (siehe unten, Datenexport)

Zur einfachen Abfrage von Daten soll eine Abfragewerkzeug in der Weboberfläche des FIS die Abfrage von Messwerten mit zeitlicher und räumlicher Einschränkung sowie Einschränkung nach Parametern, Messstellen oder deren Eigenschaften möglich sein.

Die Ergebnisse sollen mit folgenden Visualisierungsmethoden aufbereitet werden:

- ▶ Kurvendiagramm für zeitliche Verläufe von Messwerten mit einer Kurve pro Parameter an einer Messstelle oder über mehrere Messstellen aggregiert
- ▶ Kurvendiagramm für zeitliche Verläufe von Messwerten mit einer Kurve pro Messstelle oder Grundwasserkörper für einen Parameter
- ▶ Histogramme, Boxplots oder andere Visualisierungsmethoden der deskriptiven Statistik für die Verteilung von Messwerten eines Parameters an einer Messstelle oder über mehrere Messstellen aggregiert, inklusive Darstellung von Extrema, Perzentilen, Interquartilsabstand, z-Score etc.
- ▶ Zeitstrahl der Messzeitpunkte an einer oder mehreren Messstellen

Bei Auswertungen soll sowohl grundsätzlich als im Einzelfall abweichend entschieden werden können, wie Messwerte, die unterhalb einer Bestimmungsgrenze liegen, verrechnet werden (Optionen sind: als Nullwert, als halber Bestimmungsgrenzenwert, diese Werte ausschließen).

Nach Möglichkeit sollen einzelne Diagramme auch über die Schnellauswahl einer Messstelle außerhalb des Abfragewerkzeug gezeigt werden.

E.6.2 6.2 Berechnungen

Sämtliche in den Auswertungsroutinen notwendige Aggregationen sollen automatisch bei der Auswertung vorgenommen werden.

E.7 Datenexport

Messwertdaten sollen zum Zweck der weiteren Analyse mit externer Analysesoftware exportierbar sein.

Die zu exportierende Datenmenge soll dabei mit verschiedenen Filtermethoden zeitlich, nach Parameter(n) und nach Messstellen und verschiedenen Eigenschaften eingeschränkt werden können.

Die Daten sollen im Excel, CSV, JSON und einem SQL-Datenbankdump-Format exportiert werden. Der Export soll über die Weboberfläche oder mittels eines programmatischen Zugriffs über eine HTTP-API erfolgen können.

E.8 Berichterstattung

Für das WISE-6 water quality reporting an die Europäische Umweltagentur (EUA) müssen in jährlichem Turnus Messwerte des vorvergangenen Kalenderjahres berichtet werden.

Der Bericht wird dazu in einem definierten Excel- oder XML-Format auf einer EUA-Webseite hochgeladen.

Dafür soll eine Funktion zur Erzeugung und zum Download einer geeigneter XML-Datei implementiert werden, die analog zu den Anforderungen bereits im FIS Grundwasser existiert und mit geringen Anpassungen weitergenutzt werden kann.

Für die Berichterstellung sind folgende Informationen nötig, die über die Weboberfläche konfiguriert werden sollen:

- ▶ Umfang der zu berichtenden Parameter
- ▶ Für jeden zu berichtenden Parameter:
 - Code des Parameters, der von WISE-6 zur Identifikation verwendet wird.
 - Einheit, in der Messwerte dieses Parameters berichtet werden sollen
 - Bezeichnung dieser Einheit, die von WISE-6 zur eindeutigen Identifikation der Einheit verwendet wird.

Beim Upload zur EUA-Webseite findet eine Reihe von dokumentierten QS-Prüfungen statt, die im FIS bei der Berichterstellung ebenso geprüft werden sollen, sofern sie nicht schon Teil der Import-QS sind.

Für das Global Environment Monitoring System des Umweltprogramms der Vereinten Nationen (GEMS/Water) findet eine regelmäßige Berichterstattung der Messwerte statt. Hierfür liegt ein R-Skript vor, das vom FG erstellt und genutzt wird, um die Messwerte jahresweise abzufragen, zu filtern und zu transformieren und daraus CSV-Dateien in einem definierten Format zu generieren. Eine Funktion zur Erzeugung und zum Download dieser CSV-Datei soll implementiert werden.

Für diese Berichterstellung sind folgende Informationen nötig, die über die Weboberfläche konfiguriert werden sollen:

- ▶ Umfang der zu berichtenden Parameter

F MUDAB Bestandsaufnahme

F.1 Auswertung bereitgestellter Materialien

Vom Auftraggeber wurden folgende Unterlagen zur Verfügung gestellt:

- ▶ Oracle Datenbankexport vom 2022-01-31
- ▶ MUDAB_CODELISTEN.pdf
- ▶ MUDAB Datenbank.docx
- ▶ MUDAB_Datenmodell.zip (“Templates”) mit folgenden Dateien:
 - – biot_hl.pdf
 - – biot_msmnt.pdf
 - – biot_sample.pdf
 - – blg_msmnt.pdf
 - – blg_sample.pdf
 - – blg_trhl.pdf
 - – dataset.pdf
 - – mudab_projectstation.pdf
 - – parameter.pdf
 - – sediment_msmnt.pdf
 - – sediment_sample.pdf
 - – station.pdf
 - – stat_msmnt.pdf
 - – water_msmnt.pdf
 - – water_sample.pdf

- ▶ MUDAB_Jahr_Institution_Schiffcode_Datum_Kompartimente.xlsx

Die für die Beantwortung der Eingangsfragen wichtigsten Unterlagen sind dabei die Beschreibung „MUDAB Datenbank.docx“ mit den Produkten sowie den eingesetzten Werkzeugen und der Oracle Datenbankexport.

F.2 Oracle-Datenbank MUDAB

Die MUDAB wurde im Jahr 2012 vom BSH Hamburg übernommen und musste vollständig überarbeitet werden.

Die alten Daten und Modellteile sind noch in der Datenbank (Oracle MUDAB) enthalten. In der Tabelle WATER_MSMNT finden sich in der Spalte SAMPLING_DATE_WS Werte von 19860102 bis 20210315. Es hat folglich eine Altdatenübertragung stattgefunden.

Es sind (Tabelle CODE_OLD) Tabellen Bestandteil der MUDAB, die nicht mehr aktiv genutzt werden. Ein kompletter Satz alter Tabellen von vor der Überarbeitung konnte jedoch nicht identifiziert werden.

Die Oracle-Datenbank umfasst 109 Tabellen, von denen ca. 60 Codelisten sind. Als Namen wurden weitgehend Abkürzungen verwendet, die sich nur zum Teil erschließen. Daher ist eine eingehende Analyse und technische Dokumentation notwendig. Einige Kürzel (z. B. "AGDET" für age detection und "MSMT" für measurement) erschließen sich aus dem Inhalt oder sind einer PDF-Sammlung "Datenmodell" (im folgenden Text: Templates) mit 15 Templates zu entnehmen. Eine Korrelation der Templates mit Tabellen oder anderen Teilen des Datenmodells ist nicht unmittelbar erkennbar. Trotzdem sind diese für die Erlangung eines Verständnisses des Datenmodells bedeutsam, da sie Rückschlüsse auf die Semantik von Abkürzungen in den Tabellennamen zulassen.

Wesentliche Tabellen mit Messwerten, Probenahmen bzw. Messungen, Messstellen und Parametern lassen sich schnell identifizieren. Da die Oracle-MUDAB bei vielen Tabellen, an denen Stichproben bzgl. der mit Hilfswerkzeugen extrahierbaren Tabellenbeschreibungen gemacht wurden, sowohl Beschreibungen der Spalten als auch Verknüpfungsinformationen enthält, ist die Erstellung einer technischen Dokumentation in der für eine Überführung nach FIS Wasser-Boden notwendigen Tiefe möglich. Zu zahlreichen Listen, nach Stichproben voraussichtlich zu allen mit der Zeichenfolge „ICESCL_“ beginnenden, existieren im Datenmodell keine Verzeigerungen.

Bei der Untersuchung der Wertetabellen wurden Zeichenketten entdeckt, deren Inhalt derzeit nicht interpretierbar ist. Dies betrifft gerade die wichtigen Wertetabellen, MUDAB-Tabelle WATER_MSMNT und MUDAB-Tabelle PARAMETER. Letztere enthält aller Voraussicht nach (sie umfasst mehr als 5 Millionen Zeilen) ebenfalls Werte und nicht Parameter. Es besteht die Möglichkeit, dass die Tabelle PARAMETER auch beschreibende Informationen zu den Messwerten in WATER_MSMT enthält. Jedoch sollte dann eine Referenz von WATER_MSMT auf PARAMETER existieren, was nicht der Fall ist.

F.2.1 Messwerte

Tabelle WATER_MSMNT enthält 6.3 Mio Messwerte. Um welche Messgröße mit welcher Einheit es sich jeweils handelt, ist leider nicht der Spalte PARAMETERID_PM zu entnehmen, wie zunächst angenommen wurde.

Bei der Sichtung der hier beschriebenen Messwerte wurden Proben jüngeren Datums gefunden, weshalb es sich hier um einen aktiv genutzten Teil der MUDAB handeln muss, der von BfG neu erstellt wurde.

Leider existiert keine Referenz auf einen Parameter, wie er in anderen Umweltkompartimenten unter dem Begriff "Parameter" (oder Messgröße) verstanden wird. Die erwähnte Spalte PARAMETERID_PM, die folgende Beschreibung trägt, existiert jedoch:

parameterID: Kennung des Parameters

Im BLMP automatisch befüllt durch Eingabe der Schlüsselwerte - Verkettung folgender Inhalte:

- Inhalt "Datenart",

- Inhalt "Disziplin",
- Inhalt "Parametercode",
- Leerzeichen,
- Inhalt "Messmethode",
- Inhalt "phys. Behandlung",
- Inhalt "chem. Behandlung".

Für eine Integration der MUDAB in das FIS Wasser-Boden muss die Korrelation der ICES-Code-listen zu den hier genannten Inhalten hergestellt werden. Der Stichprobenwert HCDOXYO-XENONE~NFNON ml/l IFOG und auch weitere Stichprobenwerte sind nicht mit der Beschreibung in Übereinstimmung zu bringen, weshalb die Beschreibung als wenig hilfreich angesehen werden muss.

Es wird hier Datenforensik notwendig sein, um eine Interpretation der Inhalte zu ermöglichen. Es besteht eine gewisse Wahrscheinlichkeit, dass für die Interpretation die Hilfe der BfG notwendig wird.

Es existiert auch keine IT-technische Referenz auf eine Projektstation (i.S. der FIS-WaBo Messstelle) oder eine Messstelle (i.S.v. Koordinaten). Hingegen sind Datum und Uhrzeit der Probenahme sowie ein numerischer Wert, der Messwert, vorhanden.

Die Zusammenführung mehrerer Werte in einer Zeichenkettenspalte stellt ein erhebliches Risiko für die Übertragung der Daten in die Datenbank des FIS Wasser-Boden dar und wird signifikante Aufwände nach sich ziehen. Der Grund dafür ist, dass die hier in einer Spalte zusammengeführten Inhalte nicht wie z. B. bei einem CSV-Format mit Trennzeichen so versehen sind, dass wenn ein Wert nicht angegeben wurde, das Trennzeichen sich wiederholt. Somit ist eine klare Zuordnung der durch Leerzeichen getrennten Zeichenketten zu einem Attribut nicht möglich bzw. benötigt analytische Arbeit, um Teile oder möglichst alle Inhalte interpretieren zu können. Da es sich um mehrere Millionen Zeilen handelt, ist eine manuelle Bearbeitung unrealistisch.

F.2.2 Parameter

Die Tabelle Parameter enthält ca. 5 Mio. Zeilen. Bei deren Inhalt handelt es sich folglich nicht um Parameter, sondern um Werte. Die Beschreibung der Tabelle identifiziert zwei Spalten, die den Beginn und das Ende der Beprobung beschreiben sollten. Stichproben ließen Probenahmen der Jahre 2012 bis 2014 erkennen. Ein Großteil dieser Werte (4,98 Millionen) besitzt jedoch kein Probenahmeintervall. Lediglich 81336 Zeilen besitzen einen expliziten Datumsbezug, der von 19920101 bis 20211104 reicht.

Die Parameter besitzen einige beschreibende Attribute, die mit unterschiedlichen Drei-Buchstaben-Kürzeln gefüllt sind. Auch hier sind jedoch Spaltenbeschreibungen und Referenzen auf Listentabellen zu finden, so dass sich mit ausreichend Zeit das Datenmodell extrahieren und ein Verständnis über die Zusammenhänge herstellen lässt.

Worum es sich bei dem in der Tabelle Parameter abgelegten Wert handelt, ist wiederum in einer Kennung PARAMETERID_PM hinterlegt, die hier wie folgt beschrieben ist:

parameterID: Kennung des Parameters

Im BLMP automatisch befüllt durch Eingabe der Schlüsselwerte - Verkettung folgender Inhalte:

- Inhalt "Datenart"
- Inhalt "Disziplin"
- Inhalt "Parametercode"
- Leerzeichen
- Inhalt "Messmethode"
- Inhalt "phys. Behandlung"
- Inhalt "chem. Behandlung"
- Leerzeichen
- Inhalt "Einheit"
- Inhalt "Matrix (bei Sediment)"
- Leerzeichen
- Inhalt "Analyse-Labor"
- Leerzeichen
- Inhalt "interne QS / Nachweisgrenze"
- Leerzeichen
- Inhalt "interne QS / Bestimmungsgrenze"

Die Stichprobe 9 135852 HCAMON 133 SPECNFNON umol/l BSHG lässt wiederum eine Einheit und vrs. ein Laborkürzel erkennen. Der Präfix entspricht bzgl. der Anzahl der Zeichengruppen nicht der Beschreibung. Eine Interpretation der Daten kann eventuell mit IT-forensischen Mitteln erfolgen, was nichts mit meeresfachlichen Kenntnissen zu tun hat, sondern damit, IT-seitig herauszufinden, wofür eine 9 bzw. eine 133, oder in anderen Beispielen 6, 17, 13, 145 oder 2828 jeweils steht, denn eine Zuordnung zu einer Spalte im herkömmlichen Sinne (semantisch klares Attribut) ist durch die fehlende Struktur in der Zeichenkette nicht ohne verschiedene Betrachtungen (i.S.v. welche Liste könnte gemeint sein und hat diese Liste Elemente, deren ID mit der Zahl übereinstimmen; Ist es sicher keine andere Liste, die gemeint ist, die ebenfalls Elemente mit dieser ID enthält) möglich.

Es besteht eine gewisse Wahrscheinlichkeit, dass die Hilfe der BfG oder die Übersendung der tatsächlichen Interpretationsvorschrift, falls eine solche existiert, notwendig wird. Eventuell können einzelne Werte über das bestehende Portal ausgegeben werden und sofern dort die Teile der Zeichenketten einzeln klar benannt sind, kann darüber abgeleitet werden, wie die Zeichenketten zu interpretieren sind.

F.3 Templates

Die zur Verfügung gestellten Daten enthalten einen ZIP-Container MUDAB_Datenmodell, der mit dem in der Oracle-Datenbank hinterlegten Datenmodell keine offensichtliche

Überschneidungsmenge bildet. Er enthält 15 PDF-Dateien mit tabellarischen Informationen zu jeweils einer Entität, die durch die Werte „Template Name“, „Name“, „Geometry type“ gekennzeichnet sind (z. B. „Biota-Hol“, „BiotaHaul“, „None“).

Die Entitäten besitzen nummerierte Attribute, wobei eine Nummer immer auf dasselbe Attribut (Stichprobenvergleich mit anderen PDF-Tabellen) verweist. Hier finden sich Attributbeschreibungen, Datentypen, eine Angabe zu Pflicht/Optional und in einigen Fällen Verweise auf MUDAB-Codelisten, GML oder Excelformate („Excel: HH:MM“).

Der Einsatzzweck dieser Templates ist derzeit nicht klar und sollte im Zuge der Übertragung der MUDAB in das FIS Wasser-Boden geklärt werden.

F.4 Webpräsenz und deren Berichtsformate

Im Rahmen der Untersuchungen zur Abschätzung der Aufwände zur Integration der MUDAB in das FIS Wasser-Boden wurde auch die MUDAB Webpräsenz untersucht. Dabei lag besonderes Interesse auf den Berichtsformaten und den Webservices, da alle Kommunikationsfähigkeiten der MUDAB erhalten bleiben sollen.

- ▶ Auf der Webseite MUDAB, Kategorie Datenfluss, Unterkategorie Berichtsformate²⁹ gibt es einen Link³⁰ zu einer
 - Excel-Schnittstelle mit Vorlage und Schnittstellenbeschreibung. Nach Erstellung eines Zugangs ist keinerlei Information, insbesondere weder eine Excel-Vorlage noch eine Schnittstellenbeschreibung, erreichbar. Der Nutzer wird i.d.R. sofort ausgeloggt oder gelangt auf inhaltsfreie Bereiche:
 - „Wo darf ich was?\": Der Nutzende wird ausgeloggt
 - „Zurück zum Eintrag: Datamodel\": Der Nutzende wird ausgeloggt
 - „Bearbeitete Einträge: (Es wurden keine Einträge gefunden, die Sie bearbeitet haben) Zurück\"; Zurück-Knopf des Browsers:
 - Man gelangt auf eine Seite mit zahlreichen, nicht-MUDAB-bezogenen Informationen. Der Eintrag „Berichtsportal / Daten-Upload“ kann aufgeklappt werden. Diese enthält jedoch keinen Inhalt, sondern eine leere Fläche. Die Untersuchung mit Browser Developer Tools zeigt, dass dort tatsächlich nichts ist.
 - Der Eintrag „Berichtsnavigatoren / MUDAB“ ist ohne Funktion. Die Untersuchung mit Browser Developer Tools zeigt: Es ist ein Bild im .png-Format, jedoch ohne jegliche aktive Funktion.
 - „Benutzereinstellungen\": Hier sind Passwort- und Adressänderungen sowie Festlegung einer Startseite möglich. Bei der Auswahl der Startseite kann jedoch keine Seite festgelegt werden, die MUDAB-spezifisch wäre.
 - „Abmelden\": Dies funktioniert.

²⁹ <https://geoportal.bafg.de/MUDABAnwendung/>

³⁰ <http://www.wasserblick.net/servlet/is/143388>

- Bei der Untersuchung wurden mehrere Browser probiert. JavaScript und Cookies waren jeweils erlaubt, um dies als Fehlerquelle auszuschließen.
- Unterkategorie Berichtsformate, Link³¹ auf Codelisten: Hier ist dasselbe zu beobachten. Meist wird der Nutzende ausgeloggt. An anderer Stelle erhält man auch nach längerer Suche nicht die angekündigte Information.
- Ebenfalls Unterkategorie Berichtsformate, Link XML Schemata: Es gibt hier keine nutzbare Information.
- Ebenfalls Unterkategorie Berichtsformate, Link Informationsarchiv: Es gibt hier keine nutzbare Information.
- Auf der Webseite MUDAB, Kategorie Datenfluss, Unterkategorie Webservices gibt es Links auf zwei Capability-Beschreibungen von Web-Map-Services von Projekt- und Messstationen, dito Web-Feature-Services und einen SOS-(SensorObservation-)Service. Es erscheinen jeweils Beschreibungstexte. Auf den Web-Feature-Services lieferten Capability-Abfragen Daten, ebenso DescribeFeatureType-Abfragen und GetFeature-Abfragen, z. B. auf MUDAB_MUDAB_Projektstation:Projektstationen³². Es wird davon ausgegangen, dass auch die Web-Map-Services und der SOS-(SensorObservation-)Service funktionsfähig sind. Die Abfrage von Messwerten unter Nutzung von Services ist nicht vorgesehen. Eine Beschreibung zur intendierten Nutzung der Services ist nicht hinterlegt.
- Nachfolgend wurde der Zugang von der BfG der Gruppe „Ingenieurbüros (Any/ Gäste/ Ingenieurbüros)“ zugeordnet (Information per Email), woraufhin der Login nicht mehr funktioniert.
- Die Anforderung eines neuen Passworts funktioniert. Dieses wird unmittelbar per Email zugestellt.
- Die Informationen zur Excel-Schnittstelle und die Schnittstellenbeschreibung sind auf diesem Weg nicht zugänglich.
- Das oben skizzierte Verhalten hat sich auch nach Bestätigung per Email, dass der Zugang über die erforderlichen Rechte verfügt, nicht geändert.

F.5 MUDAB Datenimporte

Aus Interviews, u.a. mit der Betreiberin der MUDAB, konnten folgende Informationen gewonnen werden:

- Es ist ständige Pflege an den FME-QS-Prüfungen („Plausibilisierungen“) notwendig, da sich die Anforderungen der Zielsysteme ständig ändern.
 - Dies betrifft abhängige Pflichtfelder, Datumsvergleiche, Listenfeldprüfungen und Prüfungen gegen konstante Werte³³.

³¹ <http://www.wasserblick.net/servlet/is/143386>

³² http://geoportal.bafg.de/arcgis/services/MUDAB/MUDAB_Projektstation/MapServer/WFSServer?SERVICE=WFS&VERSION=1.1.0&REQUEST=GetCapabilities

³³ Das ist alles recht beherrschbar, kann jedoch organisatorisch/pflegetechnisch recht aufwändig sein, insbesondere, wenn die mit den Aufgaben befassten Personen nicht in der umweltfachlichen Thematik versiert sind.

- Es sind mehrere Hundert solcher Prüfungen im ETL-Werkzeug FME für die MUDAB hinterlegt.
- Prinzipiell ist die Vorgehensweise für den Datenimport der MUDAB, der des FIS Wasser-Boden sehr ähnlich: Erste Prüfungen beim Import, weitergehende Prüfungen nach dem Import nach Oracle. Es wird ein klassischer Rollback-Mechanismus genutzt, um den gesamten Datensatz wieder zu entfernen, sowie ein einziger Fehler auftritt. Dieses Vorgehen ist sehr strikt und es ist nicht nutzerfreundlich, beim ersten Fehler abubrechen.
- ▶ Es liegen UML-Modelle sowohl der gesamten Datenstruktur als auch der Prüfmechanismen vor. Diese werden benötigt, jedoch wurde die Übermittlung mit Verweis auf Sicherheitsinteressen verweigert. Stattdessen wurden UML-Grafiken übermittelt, die hilfreich sind, jedoch nicht den Informationsgehalt der UML-Modelldateien aufweisen.
- ▶ Die MUDAB-Produkte sind mindestens die Bereitstellung von Daten für:
 - Datenübertragungen an den ICES (verschiedenste Datensätze) als Datenbank für die Meereskonventionen
 - Nitratrichtlinie sowie
 - die Datenübertragung an die Meereskonventionen OSPAR und HELCOM und
 - die Europäische Umweltagentur im Zuge der SoE-Berichterstattung
 - Datenlieferungen als Antwort auf regelmäßig erfolgende, sogenannte DataCalls durch OSPAR und HELCOM, bei denen spezifische Datenlieferungen über den etablierten Prozess hinaus gefordert werden.

Folglich sind dies derzeit vier Reportings³⁴ und die Übermittlung von Abfrageergebnissen zu variierenden Fragestellungen.

- ▶ Die Meeresstrategierahmenrichtlinie wird nicht an den ICES berichtet, sondern an das Reportnet der EUA und nicht direkt aus der MUDAB. Die Berichterstattung erfolgt über die BLANO Geschäftsstelle und den WasserBlick.
- ▶ Es existiert vrs. analog zur Empfehlung im Unterkapitel „Importschnittstellen“ ein Produkt Pflegeportal für Stationen
- ▶ Es existiert ein Produkt Reportgenerator basierend auf „BIRT“ (Eclipse Foundation Business Intelligence and Reporting Tools):
 - Die/der Nutzende bekommt hier Tabellendaten auf zwei Ebenen in einer Darstellung mit vielen Daten auf engem Raum präsentiert. Das entspricht nicht heutigen Designrichtlinien, soll jedoch nach Rücksprache mit dem UBA so beibehalten werden, da es in dieser Form sehr nützlich ist.
 - BIRT ist sehr leistungsfähig und es wäre vergleichsweise aufwändig, ein entsprechendes Werkzeug neu zu implementieren. Der DataCube hat dort Lösungen, die zeitgemäß sind und sich an aktuellen Designrichtlinien orientieren.

³⁴ Unter Reporting wird hier die Übermittlung von Datensätzen über den Zustand eines Umweltkompartiments während eines definierten Beobachtungszeitraums an Empfängerorganisationen verstanden, die teils aufgrund von EU-Richtlinien, -Verordnungen und oft auch anderer, staatlich bindender Vereinbarungen durchzuführen sind.

► Weitere Produkte:

- ArcGIS Karte
- ArcGIS INSIGHTS for ArcGIS: Das Produkt orientiert sich am Leitfaden Nitratbericht. Insofern soll und kann es voraussichtlich nicht geändert werden, sondern soll unverändert Bestand haben.

F.6 Außendarstellung MUDAB derzeit

Es ist im Rahmen der MUDAB eine Sammlung an Darstellungen entstanden, die nicht wirklich mit einander verbunden sind. Dies geht besser, vor allem homogener, ist aber auf jeden Fall ein ernst zu nehmender Aufwand.

Hier ist sowohl eine Kuratierung der Inhalte angeraten als auch die Entwicklung eines Präsentationskonzepts, wenn MUDAB eine neue, eigenständige Webpräsenz erhalten soll. Hier könnten Visualisierungs- und Datenbereitstellungsoffensiven wie DataCube wertvolle Vorlagen und Vorgehensweisen liefern, in dem sie aufzeigen, wie geschmeidige, nutzerzentrierte Darstellungsformen für andere, strukturell ähnlich komplexe Umweltkompartimente aussehen. Die beste Lösung wäre sicherlich, wenn sich die Visualisierungsexperten der MUDAB annähmen und die Inhalte mit FIS Wasser-Boden für die Visualisierung bereitgestellt würden.

F.7 Untersuchungen mit der Zielstellung einer Datenübernahme

Zahlreiche MUDAB-Listen finden sich in einer einzigen Tabelle MUDAB_CODELISTEN. Daneben existieren jedoch noch weitere 60 Codelisten, die als explizite Tabellen modelliert sind. Z.B. befindet sich die globale Liste aller Arten in einer separaten Tabelle WORMS_SPECIES mit immerhin mehr als 2600 Einträgen.

Zur Dokumentation der MUDAB wurde vom UBA die von der BfG bereit gestellte Datei MUDAB_CODELISTEN (tabellenartiger PDF-Text) an den Verfahrenshersteller übermittelt. Auffällig war dabei, dass obwohl die Bereitstellung und der Import der MUDAB zeitlich in der Nähe der Übermittlung der Dokumentation lag, bereits eine Abweichung in der Liste ICESCL_AGDET festgestellt wurde. Die Dokumentation vom 2023-09-25 enthielt hier bereits einen zusätzlichen Listeneintrag V — Vertebra.

Die Bedeutung explizit referenzierter Listen kann aus der Dokumentation des jeweiligen Attributs erschlossen werden. Jedoch ist die Bedeutung der zahlreichen Listen, die in der Tabelle MUDAB_CODELISTEN gemeinsam abgelegt sind, unklar und nicht dokumentiert. Da jedoch der Ursprung der Listen in Tabelle MUDAB_CODELISTEN das ICES ist, kann aller Voraussicht nach sowohl ein aussagekräftiger Name als auch eine ebenso aussagekräftige Bezeichnung mit Hilfe von durch ICES bereit gestellte Dokumentationen ermittelt werden.

Insgesamt bleibt das Listenkonzept der MUDAB noch unklar. Eventuell ist die Tabelle MUDAB_CODELISTEN eine Reminiszenz der alten BSH-MUDAB und die (derzeit größtenteils aus anderen Tabellen referenzierten und) explizit als Tabellen modellierten Listen sind der aktuelle Ansatz — das wäre zumindest wünschenswert.