



seit 1558

From BExIS I To BExIS 2

Michael Owonibi, Andreas Ostrowski, Eleonora Petzold, Ken Prager,
Pawandeep Kaur, Birgitta König-Ries



Funded by

DFG

From BExIS 1 To BExIS 2

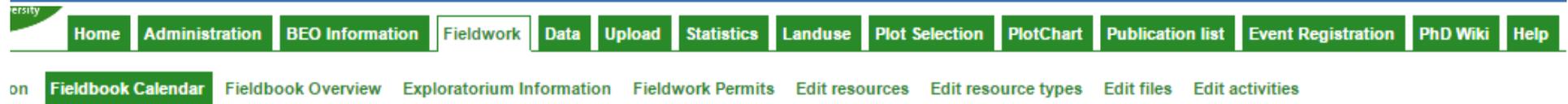


- BExIS 2 developed to make BExIS 1
 - Generic RDM tool
 - Modular
 - Extendable
 - Adaptable by being configurable
- BExIS 2 re-implemented the core functionalities of BExIS 1
- Fundamental changes
 - Functionalities
 - Data

Functionalities Differences



seit 1558



- Core functionality developed (Data, Upload & Administration)
- Functionalities not include
 - Fieldbook
 - Publications list management tool
 - Information page
 - Tools
 - Paper proposal, Landuse feedback, Climate data management, Species extraction, Darwin Core Archive generation, DQM reporting, Alumni management, Dataset overview generation, Custom dataset validation , Land owners tool etc



Functionalities Differences

- Main functionalities that may be included in other forms

- Plotchart
- Plot selection tool
- Statistics tool

Replaced by webGIS tool in BExIS 2?

Replaced by R Analysis tool in BExIS 2?

The screenshot shows a webGIS interface with a map of Europe. On the left, there is a 'Task center - Zoom to' panel with a list of locations: Germany, Hainich-Dün, Schorfheide-Chorin, and Schwäbische Alb. Below this are 'Tasks', 'Zoom to', 'Select by dataset/variables', 'Results', and 'Layers' panels. The 'Layers' panel shows 'All plots' (unchecked), 'EPs' (checked), 'EP' (red square), and 'VIP' (blue square). At the bottom left, there is a link to 'A short manual (PDF ~2.7MB)'. The map shows major cities and countries across Europe, with a search bar at the top left and navigation controls at the top right.

The screenshot shows the 'PlotChart' tool interface. It has a 'Select PLOT' section with a dropdown menu set to 'AEG1 VIP' and checkboxes for 'Include former subplots' and 'Draw beyond plot borders'. Below this is a table titled 'Edit existing/ Insert new Objects'.

NAME	GEOMETRY	TYPE	COORDINATES	COLOR	COORD_TYPE	REFERENCE_POINT	NOTES	Compute Polar
Arthropods I	rectangle		6,21,7,22	green	xy			Edit Delete
Arthropods I	rectangle		10,21,11,22	green	xy			Edit Delete
Arthropods I	rectangle		10,17,11,18	green	xy			Edit Delete
BELOW	rectangle		-6,14,5,21	red	xy		2014-2017 Auspflanzung Phytometer	Edit Delete

At the bottom of the table, there are input fields for 'rectangle', 'xy', and 'New'.

The screenshot shows the 'Basistemplate AEG1' tool interface. It features a plot visualization with a grid and several colored rectangles (red, yellow, cyan, purple) representing different plots. To the right of the plot is a 'Compute Polar Coordinates from Reference Point' section with a dropdown menu for 'Ref. Point' set to 'Core_11(-2,2)'. There are input fields for 'x', 'y', 'Target Rectangle Polar Coords (angle in cos)', and 'Nearest Point'. A 'Compute' button is at the bottom. Below the interface, there are several links for documentation: 'Overview PlotChart', 'Overview Coordinates and Geometry Types', 'How To Edit/Add Rectangle', 'How To Edit/Add Linestring', 'How To Edit/Add Error Rect', 'How To Edit/Add Circle', 'How To Edit/Add Polygon', and 'How To Compute Polar coordinates'.

Dataset Concept Differences



- Change To Dataset Concept
 - Concepts in BExIS 1 – Dataset, Primary Dataset, Metadata, Metadata schema, Variable, Matrix data
 - Concepts in BExIS 2 - Dataset , Primary data, Metadata, Primary Data Struture, Metadata Structure, Variable, Attribute, Units, Research plan
- Change To Security Concept
 - Concepts in BExIS 1 -
 - Data Rights
 - Roles (Data Rights, and Feature Permission)
 - Concepts in BExIS 1 -
 - Users and Groups
 - Feature security
 - Dataset security

Dataset Model Difference



seit 1558

- Change To Data Model
- Change To Database

Data Migration



- For functionalities already (fully/partially)developed in BExIS 2
- Get data directly from BExIS DB2 Database + Auxilliary sources
- Use of API calls to load data into BExIS 2
- Progress on
 - Dataset
 - User and security

Dataset Migration



seit 1558

- Data Structure Creation
 - Create Data Structure
 - Create Variable
 - Create or Select Attribute
 - Create or Select Unit

name	typeOfVariable	units	description
sampleID	alphanumeric		sample tag used by Rillig group
exploratory	alphanumeric		exploratory
plotNumber	alphanumeric		plot number
bagID	alphanumeric		sample label
hypPC	realNumber	percent	percent of root colonized by AMF hyphae
vesPC	realNumber	percent	percent of root colonized by AMF vesicles
arbPC	realNumber	percent	percent of root colonized by AMF arbuscules
nonPC	realNumber	percent	percent of root colonized by non-AMF fungi
hypRL	realNumber	cm/g dry soil	length of root (cm/g dry soil) colonized by AMF hyphae
vesRL	realNumber	cm/g dry soil	length of root (cm/g dry soil) colonized by AMF vesicles
arbRL	realNumber	cm/g dry soil	length of root (cm/g dry soil) colonized by AMF arbuscules
nonRL	realNumber	cm/g dry soil	length of root (cm/g dry soil) colonized by non-AMF fungi

- Issues
 - No attribute concept
 - Units not plain text in BExIS 2
 - Variable type concepts does not exactly match
- Re-useability of attributes and data structure
- Unstructured data



Mapping BExIS 1 Variables To Attributes

1	datasetid	name	type	unit	description	attribute	var_unit	cov_factor	Block	remark
614	17406	Undecane	alphanumeric	%	compound	percentage			0	DataType OK
615	17406	Unknown_terpene_mz_121_93_136	alphanumeric	%	compound	percentage			0	DataType OK
616	17406	Unknown_terpene_mz_161_133_106	alphanumeric	%	compound	percentage			0	DataType OK
617	17406	Unknown_terpenoid_mz_83_55_43	alphanumeric	%	compound	percentage			0	DataType OK
618	17406	Unknown_terpenoid_mz_91_134_1	alphanumeric	%	compound	percentage			0	DataType OK
619	17406	Unknown_terpenoid_mz_93_119_6	alphanumeric	%	compound	percentage			0	DataType OK
620	17406	Unknown_terpenoid_mz_93_71_11	alphanumeric	%	compound	percentage			0	DataType OK
621	17406	Unknown_terpenoid_mz_93_82_10	alphanumeric	%	compound	percentage			0	DataType OK
622	17406	Z3Hexenol	alphanumeric	%	compound	percentage			0	DataType OK
623	17426	vertikale_Deckung_Graeser	alphanumeric	%	vertical covering grasses	percentage			0	DataType OK
624	17426	vertikale_Deckung_krautige_Pflanzen	alphanumeric	%	vertical covering herbs	percentage			0	DataType OK
625	17426	vertikale_Deckung_offener_Boden	alphanumeric	%	vertical covering open ground	percentage			0	DataType OK
626	17426	vertikale_Deckung_Wirtspflanze	alphanumeric	%	vertical covering host plant	percentage			0	DataType OK
627	17467	Sample_No	realNumber			id_int			0	? dataType
628	17467	amoA_AOA	alphanumeric	number of copies / g DW soil	gene copies for amoA AOA	quantityPerMass	1/g	10 ³	0	? dataType
629	17467	amoA_AOB	alphanumeric	number of copies / g DW soil	gene copies for amoA AOB	quantityPerMass	1/g	10 ³	0	? dataType
630	17467	nirK	alphanumeric	number of copies / g DW soil	gene copies for nirK	quantityPerMass	1/g	10 ³	0	? dataType
631	17467	nirS	alphanumeric	number of copies / g DW soil	gene copies for nirS	quantityPerMass	1/g	10 ³	0	? dataType
632	17467	nosZ	alphanumeric	number of copies / g DW soil	gene copies for nosZ	quantityPerMass	1/g	10 ³	0	? dataType
633	17467	ammonium	alphanumeric	µg N / g DW soil	ammonium concentration DW=dryweight	ratioMass	µg/g	10 ⁻⁶	0	? dataType
634	17467	Cmic	alphanumeric	µg / g DW soil	microbial biomass	ratioMass	µg/g	10 ⁻⁶	0	? dataType
635	17467	nitrite	alphanumeric	µg N / g DW soil	nitrite concentration	ratioMass	µg/g	10 ⁻⁶	0	? dataType
636	17467	WEOC	alphanumeric	µg / g DW soil	water extractable organic carbon	ratioMass	µg/g	10 ⁻⁶	0	? dataType
637	17467	WEON	alphanumeric	µg / g DW soil	water extractable organic nitrogen	ratioMass	µg/g	10 ⁻⁶	0	? dataType
638	17467	nitrate	alphanumeric	mg N / g DW soil	nitrate concentration	ratioMass	mg/g	10 ⁻³	0	? dataType
639	17469	Sample_No	realNumber			id_int			0	? dataType
640	17469	ammonium	alphanumeric	µg N / g DW soil	ammonium concentration DW=dryweight	ratioMass	µg/g	10 ⁻⁶	0	? dataType
641	17469	Cmic	alphanumeric	µg / g DW soil	microbial biomass	ratioMass	µg/g	10 ⁻⁶	0	? dataType
642	17469	nitrite	alphanumeric	µg N / g DW soil	nitrite concentration	ratioMass	µg/g	10 ⁻⁶	0	? dataType
643	17469	WEOC	alphanumeric	µg / g DW soil	water extractable organic carbon	ratioMass	µg/g	10 ⁻⁶	0	? dataType
644	17469	WEON	alphanumeric	µg / g DW soil	water extractable organic nitrogen	ratioMass	µg/g	10 ⁻⁶	0	? dataType
645	17469	nitrate	alphanumeric	mg N / g DW soil	nitrate concentration	ratioMass	mg/g	10 ⁻³	0	? dataType
646	17506	Sample_No	realNumber			id_int			0	? dataType
647	17506	TN_root	alphanumeric	%	total nitrogen content in roots	percentage			0	? dataType
648	17506	TN_shoot	alphanumeric	%	total nitrogen content in shoot	percentage			0	? dataType

- Unclear and subjective distinction between a variable and an attribute
- As generic as possible, but that lumps different unrelated concepts together – obtaining methodology
- 19,000 Variables mapped (470 datasets), Need to check dataset to understand the meaning of each variable. . .



seit 1558

Attributes List

4	Name	ShortName	Description	IsM	IsB	Ow	Con	MeasurementScale	En	Sel	DataType	Unit
50	identifier, integer number	id_int	identifier, integer number					nominal			Integer	none
51	identifier, code based	id_char	identifier, code based					nominal			String	none
52	exploratory	exploratory	name of the explotatorium					nominal			String	none
53	area	area	area					ratio, continuous			Double	m^2
54	WGS84 coordinate	wgsCoordinate	World Geodetic System 1984 coordinate					nominal			String	deg IERS
55	decimal coordinate	decCoordinate	decimal coordinate					interval, continuous			Double	deg
56	cartesian coordinates	cartCoordinate	cartesian coordinates					interval, continuous			Double	m
57	coordinate reference system	CRS	coordinate reference system					nominal			String	none
58	description	description	description					none			String	none
59	specification, unable to order	specifUnordered	specification, scalable					nominal			String	none
60	specification, able to order	specifOrdered	scaleable order, ranking					ordinal			String	none
61	specification integer, able to order	specUno_int	integer specification, scalable					nominal			Integer	none
62	specification integer, unable to order	specOrd_int	integer scaleable order, ranking					ordinal			Integer	none
63	angle	angle	measure of a change in direction or orientation					interval, continuous			Double	deg
64	angle arc	arc	arc in degree minute second (" ' ")					nominal			String	arc
65	boolean	bool	boolean value					nominal			Bool	none
66	boolean as number	bool_int	boolean as number (e.g. 0,1)					nominal			Integer	none
67	boolean as char	bool_char	boolean as character (e.g. y/n)					nominal			String	none
68	volume	volume	three dimensional extent of an object					ratio, continuous			Double	m^3
69	relative humidity	humidityRel	relative air moisture					ratio, continuous			Double	%
70	pH	pH	measure of the acidity or basicity					ordinal			Double	none (pH)
71	value	value	parameter, characteristic, feature, measurable dimensionless					nominal			Double	none
72	atmospheric pressure	airPressure	barometric pressure					ratio, continuous			Double	Pa
73	age	age	time interval of species being in years					interval, discrete			Integer	a
74	electric charge	charge	electric charge, battery capacity					ratio, continuous			Double	Ah
75	concentration per mass	concentrationMass	relationship between concentration and mass					ratio, continuous			Double	%/kg
76	length per area	ratioLengArea	relationship between length and area					ratio, continuous			Double	m/m^2
77	wind force	wind	Beaufort wind force scale					interval, discrete			Double	Bft
78	electrical resistance	resistance	Electrical resistance					ratio, continuous			Double	ohm
79	Photosynthetically Active Photon F	PPFD	Photosynthetically Active Photon Flux Density					ratio, continuous			Double	mol/(m^2 * s)
80	electrical conductivity	conductivity	Electrical conductivity					ratio, continuous			Double	S/m
81	written range	range_string	e.g. A to B or A-B etc.					nominal			String	none

- Attribute unit consistency, data type consistency with BExIS 1 variable



Unit in Variable vs Attribute

3					
4	Name	Abbreviation	Description	DimensionName	M
116	milimeter	mm	length in milimeters	length	Ur
117	milimeterPerHour	mm/h	length in milimeters per time in hours	lengthPerTime	Ur
118	squareMilimeter	mm^2	area in square milimeters	area	Ur
119	squareMilimeterPerGram	mm^2/g	area in squaremilimeters per grams	areaPerMass	Ur
120	milimolPerAreaTime	mmol/(m^2*s)	amount in milimol per area and time	amountPerAreaTime	Ur
121	molPerAreaDay	mol/(m^2*d)	amount per area and time in days	amountPerAreaTime	Ur
122	month	months	time in months	time	Ur
123	enzymeUnitPerGram	mU/g	unit for amount		
124	milivolt	mV	voltage in miliiv		
125	nanogramPerMicroliter	ng/μl	mass in nanogra		
126	nanogramPerGram	ng/g	relationship bet		
127	nanomol	nmol	amount in nanor		
128	nanomolPerGramHour	nmol/(g*h)	amount in nanor		
129	stockSolidCubeMeter	Vfm	volume in Vorra		
130	stockSolidCubeMeterPerHectare	Vfm/ha	volume in Vorra		
131	centimeterPerSecond	cm/s	velocity in centi		
132					

1	id	name	syntax
2	1	dimensionless	
3	2	length	L(1,0)M(0,0)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
4	3	mass	L(0,0)M(1,0)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
5	4	time	L(0,0)M(0,0)T(1,0)I(0,0)P(0,0)N(0,0)J(0,0)
6	5	current	L(0,0)M(0,0)T(0,0)I(1,0)P(0,0)N(0,0)J(0,0)
7	6	temperature	L(0,0)M(0,0)T(0,0)I(0,0)P(1,0)N(0,0)J(0,0)
8	7	amount	L(0,0)M(0,0)T(0,0)I(0,0)P(0,0)N(1,0)J(0,0)
9	8	luminous	L(0,0)M(0,0)T(0,0)I(0,0)P(0,0)N(0,0)J(1,0)
10	9	watt	L(2,0)M(1,0)T(0,-3)I(0,0)P(0,0)N(0,0)J(0,0)
11	10	amountPerTime	L(0,0)M(0,0)T(0,-1)I(0,0)P(0,0)N(1,0)J(0,0)
12	11	perTime	L(0,0)M(0,0)T(0,-1)I(0,0)P(0,0)N(0,0)J(0,0)
13	12	perArea	L(0,-2)M(0,0)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
14	13	perMass	L(0,0)M(0,-1)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
15	14	lengthPerTime	L(1,0)M(0,0)T(0,-1)I(0,0)P(0,0)N(0,0)J(0,0)
16	15	areaPerMass	L(2,0)M(0,-1)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
17	16	MassPerArea	L(0,-2)M(1,0)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
18	17	MassPerVolume	L(0,-3)M(1,0)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
19	18	volumePerMass	L(3,0)M(0,-1)T(0,0)I(0,0)P(0,0)N(0,0)J(0,0)
20	19	amountPerMassTime	L(0,0)M(0,-1)T(0,-1)I(0,0)P(0,0)N(1,0)J(0,0)
21	20	massPerTime	L(0,0)M(0,-1)T(0,-1)I(0,0)P(0,0)N(1,0)J(0,0)



seit 1558

Data Structure Creation

1	datasetid	name	type	unit	description	attribute	var_unit	cov_factor	Block	remark
614	17406	Undecane	alphanumeric	%	compound	percentage			0	DataType OK
615	17406	Unknown_terpene_mz_121_93_136	alphanumeric	%	compound	percentage			0	DataType OK
616	17406	Unknown_terpene_mz_161_133_106	alphanumeric	%	compound	percentage			0	DataType OK
617	17406	Unknown_terpenoid_mz_83_55_43	alphanumeric	%	compound	percentage			0	DataType OK
618	17406	Unknown_terpenoid_mz_91_134_1	alphanumeric	%	compound	percentage			0	DataType OK
619	17406	Unknown_terpenoid_mz_93_119_6	alphanumeric	%	compound	percentage			0	DataType OK
620	17406	Unknown_terpenoid_mz_93_71_11	alphanumeric	%	compound	percentage			0	DataType OK
621	17406	Unknown_terpenoid_mz_93_82_10	alphanumeric	%	compound	percentage			0	DataType OK
622	17406	Z3Hexenol	alphanumeric	%	compound	percentage			0	DataType OK
623	17426	vertikale_Deckung_Graeser	alphanumeric	%	vertical covering grasses	percentage			0	DataType OK
624	17426	vertikale_Deckung_krautige_Pflanzen	alphanumeric	%	vertical covering herbs	percentage			0	DataType OK
625	17426	vertikale_Deckung_offener_Boden	alphanumeric	%	vertical covering open ground	percentage			0	DataType OK
626	17426	vertikale_Deckung_Wirtspflanze	alphanumeric	%	vertical covering host plant	percentage			0	DataType OK
627	17467	Sample_No	realNumber			id_int			0	? dataType
628	17467	amoA_AOA	alphanumeric	number of copies / g DW soil	gene copies for amoA AOA	quantityPerMass	1/g	10^3	0	? dataType
629	17467	amoA_AOB	alphanumeric	number of copies / g DW soil	gene copies for amoA AOB	quantityPerMass	1/g	10^3	0	? dataType
630	17467	nirK	alphanumeric	number of copies / g DW soil	gene copies for nirK	quantityPerMass	1/g	10^3	0	? dataType
631	17467	nirS	alphanumeric	number of copies / g DW soil	gene copies for nirS	quantityPerMass	1/g	10^3	0	? dataType
632	17467	nosZ	alphanumeric	number of copies / g DW soil	gene copies for nosZ	quantityPerMass	1/g	10^3	0	? dataType
633	17467	ammonium	alphanumeric	µg N / g DW soil	ammonium concentration DW=dryweight	ratioMass	µg/g	10^-6	0	? dataType
634	17467	Cmic	alphanumeric	µg / g DW soil	microbial biomass	ratioMass	µg/g	10^-6	0	? dataType
635	17467	nitrite	alphanumeric	µg N / g DW soil	nitrite concentration	ratioMass	µg/g	10^-6	0	? dataType
636	17467	WEOC	alphanumeric	µg / g DW soil	water extractable organic carbon	ratioMass	µg/g	10^-6	0	? dataType

- Created based on information from Variable-Attribute mapping
- Not re-used at the moment
- 468 records with about 18,000 variables, 77 attributes, 127 units (52 attribute-and-var-units, 75 var-units) and 36 dimensions are created in BExIS 2.
 - Duration ca 1.5 hours

Primary Data Migration



- Easier migration of unstructured data
- Structured Primary Data Migration
 - For simple table
 - Matrix data not supported
 - Table with block information not yet supported (block concept needs to be mapped to some concept in BExIS 2)
- Versioning – BExIS 1 vs BExIS 2
 - Created in BExIS 2 based on timestamp information in BExIS 1
- Performance
 - Data structure creation and upload
 - Disable validation during migration

Metadata Migration



seit 1558

- Create metadata structure
- Migrate metadata data

Metadata Structure Creation



- BExIS 2 supports metadata structure creation from a XML schema file
- Edit BExIS schema
 - Remove
 - Data structure information
 - Data file parsing information
 - Where to store information removed?
- Requirements not supported
 - XML attributes
 - Security on elements of metadata
 - Client based data entry assistance e.g data owners and manager list
 - Client & server based data validation
 - Blocking edits of elements value after entry

Metadata Data Migration

- NOTE : BExIS 2 create a new XML schema file based on but totally different from the given BExIS XML schema file
- Based on Metadata mapping XML
 - Maps elements between BExIS 1 and BExIS 2 schema
 - Manually created
- Metadata created in BExIS 2 using
 - Mapping file & Metadata read from BExIS 1 database

```
▼<route>
  <source xpath="metaProfile/general/author"/>
  <destination xpath="Metadata/general/generalType/author/author"/>
</route>
▼<route>
  <source xpath="metaProfile/general/originalDatasetManager"/>
  <destination xpath="Metadata/general/generalType/originalDatasetManager/originalDatasetManager"/>
</route>
▼<route>
  <source xpath="metaProfile/general/owners/owner"/>
  <destination xpath="Metadata/general/generalType/owners/ownerType/owner/owner"/>
</route>
▼<route>
  <source xpath="metaProfile/general/notification/downloadNotification"/>
  <destination xpath="Metadata/general/generalType/notification/notificationType/downloadNotification/downloadNotificati
</route>
▼<route>
  <source xpath="metaProfile/general/projectName"/>
  <destination xpath="Metadata/general/generalType/projectName/projectName"/>
</route>
```

User Migration



- User Information functionality not yet in BExIS 2
- Username, user's name, email migrated
- Incompatible password encryption (hashing) system
 - Generate random password for each user in BExIS 2
- Feature permission
 - Mapping of concepts in BExIS 1 to 2 (to be done)
 - BExIS 2 features vs BExIS 1 features e.g. Data Collection, Research Plan and Data Dissemination
 - Roles vs Groups
 - Data Create, View, Update, Delete, Download vs Data View & Edit
 - A test case
 - Create a BExIS 2 „Group“ for all BExIS 1 „users“ and grant it „feature permission“
 - Create dataset and give owners all dataset permissions
 - Duration for 488 users ca 5 min

Open Issues



seit 1558

Metadaten

Die in BEXIS1 in den Metadaten enthaltene Datenstruktur ist in BEXIS2 nicht mehr in den metadaten enthalten.

Variablen werden in BEXIS2 in der dataStructure gehandelt.

Umsetzung von block und matrix in BEXIS2 unklar (je nach Anzahl betroffener Datensätze muss eine Implementierung im BMM abgewogen werden).

Für andere Werte aus den Metadaten, die noch nicht übertragen wurden, ist unklar ob diese noch benötigt bzw wo sie in BEXIS2 abgelegt sollen.

In BEXIS1 gibt es in der metadaten-xml xml-attribute, die für Funktionalitäten von BEXIS1 verwendet werden. Diese werden nicht in die BEXIS2-metadaten-xml übertragen. Es muss auf die Umsetzung dieser Funktionalitäten in BEXIS2 gewartet werden.

Leere metadatenknoten werden nicht angezeigt. Dies müsste in BEXIS2 noch geändert werden (David).

500 Datensätze sind gemappt (unter Vorbehalt). Der Rest steht noch aus

!!nur im Ausnahmefall neue Attribute hinzufügen!!!!!!!

Für die gemappten Datensätze gibt es zwei Listen mit behobenen bzw. noch existierenden Problemen, die noch gelöst werden müssen.

upload primary data:

Im mapping file wird der datentyp für das verwendete dataAttribute zugewiesen, dabei kann es vorkommen (insbes. bei dateTime) dass die Werte nicht konvertiert werden können.

1. Möglichkeit:

Methode zur Konvertierung entwickeln. z.B. Datum:

Zu jedem datensatz muss das Datumsformat zugeordnet werden. Ein Datum in BEXIS1 vom Typ dateTime besitzt das in den Metadaten angegebene format (als dateTime geht die Konvertierung sowieso). Bei datum-variablen vom Typ string muss das angegebene Format geprüft bzw das entsprechende Format verwendet werden.

2. Möglichkeit:

Mapping-file anpassen.

D.h. statt attribute "dateTime" das attribute "dateTime_string" verwenden.

Überarbeitung der DataAttributes:

Trennung des Attributes "dateTime" in drei Attribute "dateTime", dateTime_date" and dateTime_time (evtl. auch für dateTime_string und dateTime_int). Neben dem BEXIS2-Datentyp "dateTime" müssen zusätzlich die BEXIS2-datentypen "time" und "date" erstellt/verwendet werden

alle drei verwenden den SystemTyp DateTime, aber die Verwendung dieser Spezialisierungen ermöglicht eine bessere Rückgabe der Daten zum user (z.B. time = 8:00:00 statt 30.12.1899 8:00:00)

Dies erfordert eine Anpassung des mapping-files!!!

Alternativ ist die verwendung von patterns auf attribute bzw. insbes hier auf variablen-ebene möglich um den gleichen Effekt zu bewirken.

Prüfen, ob die Verwendung von variable-local-names die Wiederverwendung von datasets erhöht.

Idee: variabl-label wird vom attribute-name beerbt und wird nicht verändert. Der vom user verwendete Variablenname wird als variable-local-name auf Datensatz-Ebene organisiert. Für die Prüfung, ob eine datenstruktur dann gleich ist genügt der Vergleich von attribute(ID) UND unit(ID).

für meine 468 wurden 361 datenstrukturen angelegt (reusing hauptsächlich unstructured ds); für das verwenden von local-names wären es nur noch 329 datenstrukturen.

Prüfen der Notwendigkeit des hinzufügen von patterns auf variablen- und/oder attribut-ebene in das mapping-file

Primärdaten:

Im Moment: Lesen der observations je timestamp, erzeugen der tuples und schreiben bzw adden zur tuple-list.

Probleme für große Datensätze, dass der Speicher nicht ausreichen könnte, sollte nicht auftreten, da die upload-files in BEXIS1 in ihrer Größe beschränkt sind. Wenn mehrere files geuploadet wurden besitzen die observations auch verschiedene timestamps. Dies führt wiederum zu mehreren durchläufen der Schleife und damit zu verschiedenen versionen im BEXIS2.

(Bin mir aber nicht sicher, da die Daten nach dem überführen in die tuples zweimal im Speicher sind und die tuple-xml deutlich größer ist als die observation-xml.)

timestamps vereinigen:

BEXIS1-Versionen (timestamps) die nicht lange auseinander liegen sollten zusammengeführt werden.

Effekt: falls einzelne Zeilen nacheinander geuploadet oder editiert wurden.

Problem: obiger effekt des nicht auftretenden Speicherüberlaufes ist gefährdet. Zur Performance:

Diese Funktion dauert mit steigender tupleZahl in der version-table erheblich länger: GetDatasetVersionEffectiveTupleIds

Die Laufzeit der Methode ValidateRow im XmlDataReader mach (fast) die gesamte Laufzeit der Schleife über die Observations aus.

Da die Daten aus BEXIS1 bereits valide sind kann der Aufruf dieser Funktion ausgelassen werden.

Dies erfordert jedoch, dass die DatenTypen in BEXIS2 jenen aus BEXIS1 genau entsprechen müssen. Dies wird im mapping-file gehandelt.

Im Moment ist dies insbes. bei Datum nicht der Fall.!!!!!!! mapping-file muss dafür unbedingt überarbeitet werden!!!!!!!

Für große Tuple-Lists dauert EditDatasetVersion lange (ist aber wahrscheinlich nicht zu ändern)

Überarbeiten von dateTime im mappin-file

UND datenTypen vergleichen! - unklare sind in der remark mit "? dataType" gekennzeichnet.

Problem sind oft Werte wie NA oder NAN usw.

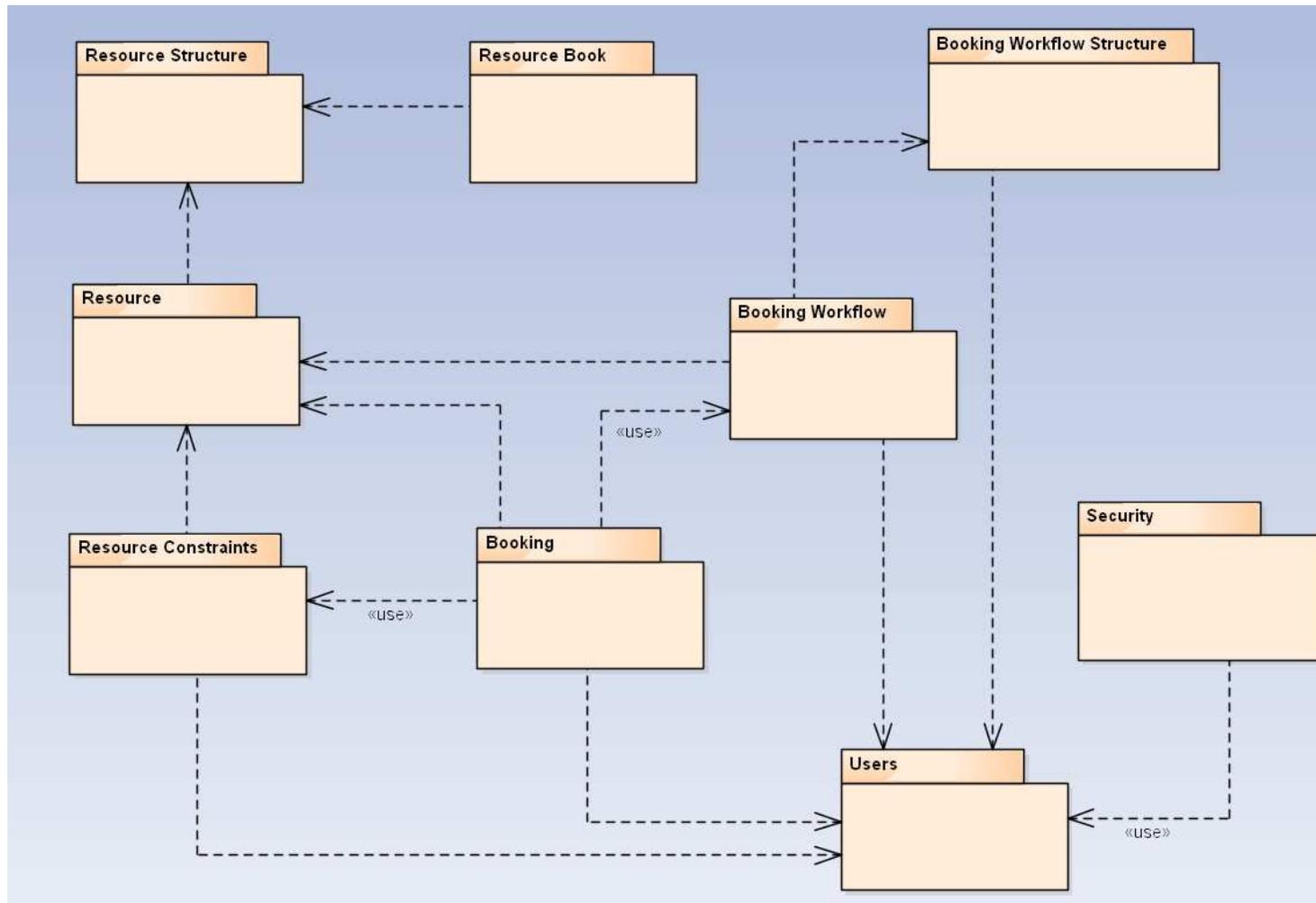
- Die originalen (vom user in b1 angelegten) variablen-units müssen noch übernommen werden (z.B. in ein extra feld)
- Denn: die units wurden für reusing und dimension-zuweisung überarbeitet; bzw können zur Zeit bei einigen (string-type-)attribute gar nicht angelegt werden.

Functionality Migration



seit 1558

- Generic, configurable Resource Book Management Functionality



Screenshots from RBM implementation



seit 1558

New Resource

Actions	Id	Name	Resource Structure Name	Description
	6	Metal Detector	Field equipment	Metal Detector
	5	Key to the flat	Residence	Key for the project house
	4	Binocular	Field equipment	Binocular
	2	Grünland-Nationalpark	Area	HEG2, HEG4, HEG7, HEG11
	1	Stiftung Schorfheide-Chorin	Area	SEW1, SEW2, SWE24, SEW33

Create new Resource

Resource Name

Resource Description

Resource Quantity

Maximum count number of shares

Resource Structure

Area

Habitat

Plotlevel

BookingTimeGranularity

Duration

LargeUnitOfTime

Book Resource

March 2015

Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31	1	2	3	4
5	6	7	8	9	10	11

Book Resource

Date Resources

StartDate

EndDate

Book Resource

Select	Id	Name	Description	Availability
<input type="checkbox"/>	6	Metal Detector	Metal Detector	<input type="button" value="Check"/>
<input type="checkbox"/>	5	Key to the flat	Key for the project house	<input type="button" value="Check"/>
<input type="checkbox"/>	4	Binocular	Binocular	<input type="button" value="Check"/>
<input type="checkbox"/>	2	Grünland-Nationalpark	HEG2, HEG4, HEG7, HEG11	<input type="button" value="Check"/>
<input type="checkbox"/>	1	Stiftung Schorfheide-Chorin	SEW1, SEW2, SWE24, SEW33	<input type="button" value="Check"/>

Publications List Migration



scit 1558

- Re-use of metadata + metadata structure functionality

List of Publications
Add Edit Publication files Synthesis Proposal List

VIEW OPTION

Overview
 Citation
 German Overview

CRITERIA

- ▶ Project
- ▶ Year
- ▶ Status
- ▶ Typeofpubs
 - Grassland
 - Forest
 - Plotlevel
 - ALB
 - SCH
 - HAI
 - Aboveground
 - Belowground
 - Field
 - Laboratory
 - Review
 - Infrastructures
 - BioticDataTaxa
 - BioticDataTypes
 - EnvironmentalDescriptors
 - ProcessesServices

PUBLICATIONS

Page size: Sort expression: Sort direction:

	ID	Projectname	Title	First Author(s)	Year	Status	Type	URI	PDF
Show Details	21	Vertebrates I	Local- and landscape-scale forest attributes differ in their impact on bird assemblages across years in forest production landscapes	Wells K.	2011	published	paper	available at www.sciencedirect.com	
Show Details	41	BEO	Implementing large-scale and long-term functional biodiversity research: The Biodiversity Exploratories	Fischer M.	2010	published	paper	available at www.sciencedirect.com	
Show Details	61	BExIS	A flexible statistics web processing service - Added value for information systems for experiment data	Heimann D.	2010	published	paper	available at journal.imbio.de	
Show Details	62	BExIS	The Biodiversity Exploratories Information System - Towards a service-oriented Framework for knowledge-based Data and Tool Integration	Heimann D.	2009	published	conference proceedings	available at www.tr32db.uni-koeln.de	
Show Details	63	BExIS	A web service based approach for integrating statistics tools into an information system for experiment data	Heimann D.	2009	published	conference proceedings	available at subs.emis.de	
			Die Verwendung von						

List of Publications **Add** Edit Publication files Synthesis Proposal List

Publication details:

Project:

First author:

Co-authors (comma separated):

Institutions (comma separated):

Email of corresponding author:

Title of the publication:

German translation of original title:

Keywords (comma separated):

Publication type:

Published in:

DOI (optional):

URL reference (optional):

Citation (as listed in a paper):

Year of publication:

Status:

Summary/Abstract (max 8000 characters):

German summary/abstract (max 8000 characters):

Exploratory objects:

Exploratory:

Schwaebische Alb:

Schorfheide-Chorin:

Hainich-Duen:

Plot level:

Habitat:

Grassland:

Forest:

Layer:

Integrating implementations in BExIS 2



- Implementations to be provided as modules
- Guidelines and best practises to be published



seit 1558

Thanks For Your Attention



Any Questions?