

Taxonomischer Service für Biologische Datenbanken

version 0.3

Florian Jansen

20. Dezember 2013

Zusammenfassung

In diesem Dokument wird der taxonomische Service für biologische (Fund-)Datenbanken beschrieben, wie er im Projekt „Rote Listen 2020 - Vorbereitung“ entwickelt wird.

1 Vorbemerkungen

Um die Funktionalitäten, den Datenfluss und die Designentscheidungen zu dokumentieren, werden in diesem Dokument Beispieldaten aus der *Festuca ovina*-Gruppe verwendet, um einzelne Prozessschritte sowie Dateninhalte und Formate zu erläutern. Die Tabellen werden bei der Dokumenterstellung automatisch erzeugt, zeigen also gleich die Machbarkeit oder mögliche Fehler des Ansatzes.

2 Zielsetzung

Den Datenfluss, wie er im Laufe des Jahres 2014 mit MultibaseCS, Deutschlandflora 2.0 und Vegetweb 2.0 erprobt und umgesetzt werden soll, exemplarisch aufzuzeigen. Mit dieser neuen Art des Umgang mit den taxonomischen Backbones biologischer Datenbanken, soll deren Inhalt besser interpretierbar werden und der Austausch zwischen verschiedenen biologischen Datenbanken vereinfacht werden.

3 Taxonomische Referenzlisten versus Checklisten

Wir können unterschiedliche taxonomische Konzepte, die wir in einer einzelnen Beobachtungsdatenbank speichern wollen/müssen in verschiedenen Arten und Weisen abspeichern. Die bisher in aller Regel verwendete versucht den unterschiedlichen Umfang von Taxa durch künstliche taxonomische Ebenen wie „s. str.“, „s. l.“ oder Aggregate in eine bestehende Sicht, respektive Hierarchie zu integrieren.

Deutlich wird dies, wenn wir uns eine bekannte Beispielgruppe aus dem Bereich der Gefäßpflanzen anschauen (Abbildung 1).

In einer typischen Referenzliste (hier die GermanSL mit der Standardsicht Wissk. & Haeupler 1998) sieht dies wie folgt aus:

<u>Name</u>	<u>Synonym zu</u>
<i>Festuca ovina</i> agg.	
<i>Festuca guestfalica</i> et ovina	
<i>Festuca ovina</i> s. str.	
<i>Festuca ovina</i> subsp. <i>duriuscula</i>	<i>Festuca brevipila</i>

Konzepte wie die Sicht von Jäger und Werner (2005) auf *Festuca ovina* (gleich *F. ovina* plus *F. guestfalica* aus Sicht von Wiskirchen u. Haeupler 1998) können nur durch Ergänzungen zum taxonomischen Namen kenntlich gemacht werden und bereiten Schwierigkeiten beim Einbinden in die taxonomische Hierarchie.

Ascherson (1864)	Festuca ovina L.			
	F. ovina ssp. vulgaris Koch		F. ovina ssp. duriuscula (L.) Ascherson	
	F. ovina ssp. vulgaris var. vulgaris	F. ovina ssp. vulgaris var tenuifolia (Sibth.) Ascherson		
Jäger & Werner (2005)	F. ovina agg.			
	F. ovina L.		F. filiformis Pouit.	F. brevipila Tracey
	F. ovina ssp. ovina	F. ovina ssp. guestfalica (Boenning ex Rchb.) Richt.		
Wisskirchen & Haeupler (1998)	F. ovina agg.			
	F. ovina L.	F. guestfalica Boenn. ex Rchb.	F. filiformis Pouit.	F. brevipila Tracey

Abbildung 1: Schema der unterschiedlichen Breite der *Festuca ovina* L. Konzepte dreier häufig gebrauchter Quellen.

In Checklisten oder Standardlisten werden Taxa hingegen ohne Überlappungen und nur mit den primären taxonomischen Rängen dargestellt. Dafür fehlt hier in aller Regel die Information über den Umfang der jeweiligen Art, der dann nur aus dem Zusammenhang bzw. im Vergleich zu einer anderen Checkliste rekonstruiert werden kann. Um alle in biologischen Funddatenbanken vorkommenden Konzepte abbilden zu können benötigen wir auf jeden Fall mehrere Klassifikationen (s. die drei Beispaltspalten in der folgenden Tabelle):

sec. Ascherson		sec. Rothmaler		sec. Wisskirchen
Festuca ovina	>	Festuca ovina	>	Festuca ovina
F. ovina ssp. vulgaris	>	F. ovina ssp. ovina		
		F. ovina ssp. guestfalica	≅	F. guestfalica
F. ovina var. tenuifolia	≅	F. filiformis	≅	F. filiformis
F. ovina ssp. duriuscula	≅	F. brevipila	≅	F. brevipila

Wenn wir Daten aus diesen unterschiedlichen Klassifikationen auswerten wollen, brauchen wir darüber hinaus Informationen über die Konzeptbeziehungen (Angaben aus der Mengenlehre wie „ist kongruent“, „ist kleiner“, „ist größer“) zwischen den Taxa dieser unterschiedlichen Klassifikationen.

Referenzlisten für biologische Datenbanken werden bisher in aller Regel in der ersten Form (s. Fig. ??) als flache Listen geführt, in die je nach Bedarf zusätzliche Sichten *eingebaut* werden, die eigentlich nicht in die taxonomische Sicht dieser Klassifikation passen, sondern hier lediglich als (pro parte) Synonyme

geführt würden, die aber benötigt werden, um (historische) Daten in der Datenbank abbilden zu können. Da dies nur durch die Einführung weiterer taxonomischer (Pseudo-)Ränge möglich ist, werden diese Taxa mit Zusatzangaben wie *sensu stricto*, *sensu lato* gekennzeichnet und können nur durch eine oder mehrere Zusatzränge zu den eigentlich taxonomisch als sinnvoll erachteten Taxa in Beziehung gesetzt werden.

Tabelle 1: Referenzlistenschema mit „integrierten“ Sichten. Die Spaltennamen hier und im Folgenden nach *vegdata* version 0.6, angelehnt an das Taxon Concept Scheme (TCS). Die Daten entstammen realen Tabellen, wie sie dann auch in den Anwendungsdatenbanken lokale Verwendung finden können.

Reference list used: FestucaClass							
TaxonUsageID	TaxonName	AUTHOR	AccordingTo	TaxonConceptID	TaxonConcept	TaxonRank	IsChildTaxonOfID
2385	Festuca ovina agg.	<NA>	Wissk. u. Haeupler 1998	2385	Festuca ovina agg.	AGG	61391
2396	Festuca ovina s. str.	L.	Wissk. u. Haeupler 1998	2396	Festuca ovina	SPE	91404
6483	Festuca guestfalica	Boenn. ex Rchb.	Wissk. u. Haeupler 1998	6483	Festuca guestfalica	SPE	91404
13541	Festuca ovina var. vulgaris	Koch	Wissk. u. Haeupler 1998	91404	Festuca guestfalica et ovina	VAR	0
13547	Festuca ovina subsp. guestfalica	Richt.	Jäger & Werner (2005)	91402	Festuca guestfalica * guestfalica	SSP	0
21820	Festuca duriuscula	auct.	Wissk. u. Haeupler 1998	26591	Festuca brevipila	SPE	0
21824	Festuca filiformis	Pourr.	Wissk. u. Haeupler 1998	21824	Festuca filiformis	SPE	91430
25921	Festuca ovina var. tenuifolia	Mert. & Koch	Wissk. u. Haeupler 1998	21824	Festuca filiformis	VAR	0
26591	Festuca brevipila	Tracey	Wissk. u. Haeupler 1998	26591	Festuca brevipila	SPE	91432
90000	Festuca ovina subsp. vulgaris	Koch	Ascherson 1864	2385	Festuca ovina agg.	SSP	0
90001	Festuca ovina subsp. duriuscula	Ascherson	Ascherson 1864	26591	Festuca brevipila	SSP	0
91404	Festuca guestfalica et ovina	-	[ad-hoc-Taxon GermanSL]	91404	Festuca guestfalica et ovina	AGG	91430
91415	Festuca ovina s. l.	L.	Jäger & Werner (2005)	91404	Festuca guestfalica et ovina	SPE	0
91418	Festuca ovina subsp. ovina	-	Jäger & Werner (2005)	2396	Festuca ovina	SSP	0

Bei großen Datensätzen und langer taxonomischer Geschichte wächst die Zahl der Pseudoränge, die benötigt werden, um alle Sichten integrieren zu können. Zum Teil müssen Taxa namentlich kombiniert werden, wie bei *Festuca guestfalica et ovina*, um ihren taxonomischen Inhalt noch deutlich machen zu können.

Alternativ dazu können die Taxonkonzepte aber auch **unter Angabe ihrer Klassifikation** nebeneinander bzw. untereinander geschrieben werden. Dies ermöglicht die saubere Trennung der Sichten, das heißt Klassifikationen. Wenn die Daten ausgewertet werden sollen (und erst dann!), müssen allerdings die Konzeptbeziehungen aller verwendeten Taxa zur jeweils gewünschten Auswertungssicht bekannt sein.¹

Tabelle 2: Kompilierte Referenzliste aus verschiedenen taxonomischen Quellen.

Classification	TaxonUsageID	TaxonName	AUTHOR	TaxonRank
Ascherson 1864	20	Festuca	L.	GAT
Ascherson 1864	21	Festuca ovina agg.	<NA>	AGG
Ascherson 1864	22	Festuca ovina	L.	SPE
Ascherson 1864	23	Festuca ovina subsp. vulgaris	Koch	SSP
Ascherson 1864	24	Festuca ovina subsp. duriuscula	Ascherson	SSP
Ascherson 1864	25	Festuca ovina var. vulgaris	<NA>	VAR
Ascherson 1864	26	Festuca ovina var. tenuifolia	Ascherson	VAR
Jaeger u. Werner 2005	10	Festuca	L.	GAT
Jaeger u. Werner 2005	11	Festuca ovina	L.	SPE
Jaeger u. Werner 2005	12	Festuca ovina subsp. ovina	<NA>	SSP
Jaeger u. Werner 2005	13	Festuca ovina subsp. guestfalica	Richt.	SSP
Jaeger u. Werner 2005	14	Festuca filiformis	Pourr.	SPE
Jaeger u. Werner 2005	15	Festuca brevipila	Tracey	SPE
Wissk. u. Haeupler 1998	1	Festuca	L.	GAT
Wissk. u. Haeupler 1998	2	Festuca ovina	L.	SPE
Wissk. u. Haeupler 1998	3	Festuca guestfalica	Boenn. ex Rchb.	SPE
Wissk. u. Haeupler 1998	4	Festuca filiformis	Pourr.	SPE
Wissk. u. Haeupler 1998	5	Festuca brevipila	Tracey	SPE
Wissk. u. Haeupler 1998	6	Bromus ovinus	(L.) Scop.	SPE
Wissk. u. Haeupler 1998	7	Avena ovina	(L.) Salisb.	SPE

¹Der Einfachheit halber werden in diesem Dokument für die ID's Ganzzahlen statt GUID's verwendet.

Tabelle 3: Hierarchische Beziehung der *Festuca ovina* Taxa aus Sicht von Wisskirchen und Haeupler 1998.

TaxonUsageID	TaxonName	AUTHOR	TaxonRank	IsChildTaxonOfID	IsChildTaxonOf
1	Festuca	L.	GAT	0	Poaceae
2	Festuca ovina	L.	SPE	1	Festuca ovina
3	Festuca guestfalica	Boenn. ex Rchb.	SPE	1	Festuca ovina
4	Festuca filiformis	Pourr.	SPE	1	Festuca ovina
5	Festuca brevipila	Tracey	SPE	1	Festuca ovina
6	Bromus ovinus	(L.) Scop.	SPE	0	<NA>
7	Avena ovina	(L.) Salisb.	SPE	0	<NA>

Tabelle 4: Konzeptbeziehungen der Nicht-Standard *Festuca ovina* Taxa zur Sicht von Wisskirchen und Haeupler 1998.

TaxonUsageID	TaxonName	Classification	Relation	ID	AcceptedName
25	Festuca ovina var. vulgaris	Ascherson 1864	<	1	Festuca
26	Festuca ovina var. tenuifolia	Ascherson 1864	=	4	Festuca filiformis
23	Festuca ovina subsp. vulgaris	Ascherson 1864	>	2	Festuca ovina
24	Festuca ovina subsp. duriuscula	Ascherson 1864	=	5	Festuca brevipila
22	Festuca ovina	Ascherson 1864	<	1	Festuca
22	Festuca ovina	Ascherson 1864	>	2	Festuca ovina
13	Festuca ovina subsp. guestfalica	Jaeger u. Werner 2005	=	3	Festuca guestfalica
12	Festuca ovina subsp. ovina	Jaeger u. Werner 2005	=	2	Festuca ovina
11	Festuca ovina	Jaeger u. Werner 2005	>	2	Festuca ovina
11	Festuca ovina	Jaeger u. Werner 2005	>	3	Festuca guestfalica

Die Konzepte können auf diese Weise stringenter, d.h. widerspruchsfreier und vollständiger abgebildet werden. Allerdings muss genau überlegt werden, welche Klassifikationen benötigt werden. Die taxonomische Hierarchie ist nun für jede Klassifikation getrennt zu betrachten (s. Tab. 3). Bei einem Wechsel der Standardsicht (= die für die Auswertung gewünschte Sicht) muss sie erneuert werden.

In dieser Hierarchie fehlen alle Taxa aus Nicht-Standardsichten. Diese sind nur durch die Abfrage der Beziehungen zwischen den Konzepten aus diesen Listen in Bezug auf die Standardsicht ersichtlich. Hierfür wird eine weitere Liste geladen, in der für alle Nicht-Standardsicht-Taxa die Beziehung(en) zu einem oder mehreren Standardsichttaxa dargestellt wird.

4 Die Eingabe biologischer Beobachtungsdaten

1. Eine Standardklassifikation auswählen
2. ggf. Filter wie Region und Vorkommenstatus auswählen
3. alle bzw. gefilterte Taxa der gewählten Klassifikation von der zentralen taxonomischen Datenbank herunterladen und als lokale Referenzliste speichern
4. Für Beobachtungen wird aus der lokalen Liste ein Taxon ausgewählt und die Beobachtung diesem Taxon (bzw. dessen GUID) zugeordnet
5. kann in der lokalen Liste kein passendes Taxon gefunden werden, wird aus der zentralen Datenbank zuerst ein Taxon aus einer anderen Region oder Klassifikation ausgewählt und der lokalen Datenbank hinzugefügt oder es wird ein vollständig neues (vorläufiges) Taxon zentral erzeugt

Als Beispiel soll uns wieder die *Festuca ovina* Gruppe dienen. Wir nutzen viele verschiedene Konzepte, um Daten in unserer Beobachtungsdatenbank abzuspeichern (s. Tab. 5). Erst in Kombination mit der

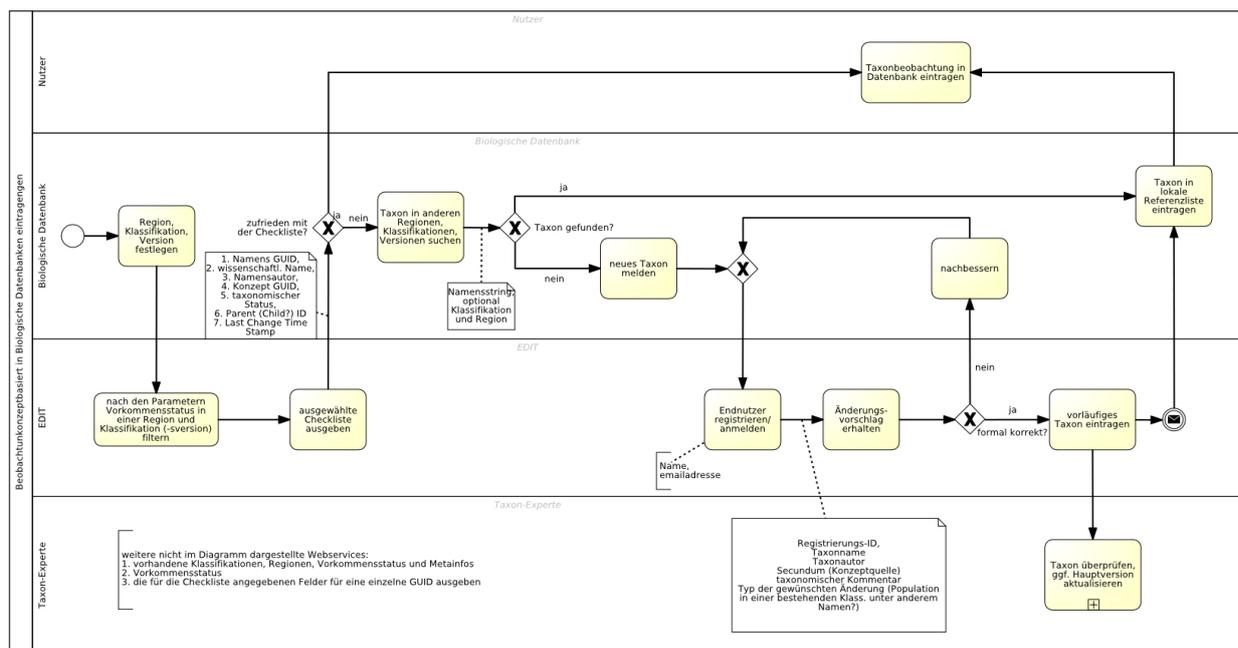


Abbildung 2: Prozessmodell für den Abruf vorhandener Klassifikation aus einem zentralen Taxonomischen Datenrepository. Zusätzlich zum in diesem Dokument skizzierten Vorgehen ist hier auch die dezentrale (vorläufige) Ergänzung von Taxa für die Nutzung in biologischen Datenbanken modelliert. Außerdem sind hier die Rollen, das heißt auch Zuständigkeiten als sogenannte Lanes (horizontale Streifen) modelliert.

zentralen taxonomischen Datenbank mit ihren Konzeptbeziehungen, bzw. den lokal zwischengespeicherten Taxonliste(n) wird inhaltlich klar, was beobachtet und abgespeichert wurde.

5 Das Zusammenführen biologischer Beobachtungs-Datenbanken

Sollen nun biologische Datenbanken zusammen geführt werden, so muss jedes Mal, egal ob konzeptübergreifende Referenzlisten benutzt wurden oder verschiedene Checklisten, mühsam geprüft werden, ob gleiche Namen auch gleiches bedeuten, d.h. welche Taxonkonzepte (welche der drei Umgrenzungen von *Festuca ovina*) gemeint sind und in welcher Beziehung diese zueinander stehen.

Um den zweiten Schritt dieser Arbeit nicht wieder machen zu müssen, sollten die Beziehungen zwischen den Konzepten verschiedener taxonomischer Sichten möglichst zentral abgespeichert werden. Im Rahmen des Rote-Lsiten 2020 Projektes werden wir für die Gefäßpflanzen Deutschlands eine taxonomische Datenbank der wichtigsten taxonomischen Quellen aus den letzten Jahrzehnten sowie ihrer Beziehungen zueinander aufbauen. Das Erstellen von Konzeptbeziehungen zwischen jeweils zwei Checklisten (mindestens eine davon mit Synonymen) kann dabei mit massiver Unterstützung von Vergleichstools der Biodiversitätsinformatik (z.B. CoTax von Gerhard Ludwig) erfolgen.

Noch besser und mittelfristig anzustreben ist der Gebrauch von Global Unique Identifiern (GUID) aus dieser zentralen Taxonomischen Datenbank in allen biologischen Datenbanken.

6 Die Auswertung von biologischen Beobachtungs-Datenbanken

Auf Datenbankebene ist nun alles klar und ein reines Datenrepository müsste gar nicht mehr wissen. Sobald wir nun aber eine Auswertung machen wollen, zum Beispiel eine Verbreitungskarte von *Festuca ovina* L. im Sinne von Wisskirchen & Haeupler 1998 darstellen, so müssen wir uns fragen, welche Beobachtungen (das

Tabelle 5: In der Biologischen Datenbank müssen Taxa aus verschiedenen Klassifikationen (=taxonomischen Quellen) gespeichert werden können.

GUID's der Beobachtungen:
 [1] 3 3 2 3 4 5 6 7 11 12 13 14 15 16 21 22 23 24 25 26
 Und ihre Bedeutungen.

TaxonUsageID	TaxonName	AUTHOR	SYNONYM	Classification	TaxonRank
3	Festuca guestfalica	Boenn. ex Rchb.	FALSE	Wissk. u. Haeupler 1998	SPE
2	Festuca ovina	L.	FALSE	Wissk. u. Haeupler 1998	SPE
4	Festuca filiformis	Pourr.	FALSE	Wissk. u. Haeupler 1998	SPE
5	Festuca brevipila	Tracey	FALSE	Wissk. u. Haeupler 1998	SPE
6	Bromus ovinus	(L.) Scop.	TRUE	Wissk. u. Haeupler 1998	SPE
7	Avena ovina	(L.) Salisb.	TRUE	Wissk. u. Haeupler 1998	SPE
11	Festuca ovina	L.	FALSE	Jaeger u. Werner 2005	SPE
12	Festuca ovina subsp. ovina	<NA>	FALSE	Jaeger u. Werner 2005	SSP
13	Festuca ovina subsp. guestfalica	Richt.	FALSE	Jaeger u. Werner 2005	SSP
14	Festuca filiformis	Pourr.	FALSE	Jaeger u. Werner 2005	SPE
15	Festuca brevipila	Tracey	FALSE	Jaeger u. Werner 2005	SPE
NA	<NA>	<NA>	<NA>	<NA>	<NA>
21	Festuca ovina agg.	<NA>	TRUE	Ascherson 1864	AGG
22	Festuca ovina	L.	FALSE	Ascherson 1864	SPE
23	Festuca ovina subsp. vulgaris	Koch	FALSE	Ascherson 1864	SSP
24	Festuca ovina subsp. duriuscula	Ascherson	FALSE	Ascherson 1864	SSP
25	Festuca ovina var. vulgaris	<NA>	FALSE	Ascherson 1864	VAR
26	Festuca ovina var. tenuifolia	Ascherson	FALSE	Ascherson 1864	VAR

heißt welche Taxon-ID's) aus unserer Datenbank, unabhängig davon aus welcher Klassifikation sie stammen, wohl zu dieser Karte dazu gehören.

Da wir für jedes Taxon aus einer Nicht-Standard Klassifikation eine Konzeptbeziehung zu einem Standardtaxon aus dem CDM abgerufen haben, ist das Vorgehen wie folgt:

1. wir fragen unsere Standardliste nach den Kindern von *Festuca ovina* L. (in unserem Falle gibt es keine)
2. für *Festuca ovina* L. und ggf. seine Kinder fragen wir nach allen alternativen Konzepten, die in unserer Konzepttabelle als kongruent oder kleiner als diese gekennzeichnet sind
3. alle Beobachtungen aus diesem Taxon-Schwarm werden aus der lokalen Beobachtungsdatenbank heraus gesucht und zusammen ausgewertet

Die folgende Mail von Andreas Müller vom 27.11.2013 geht auf die Frage ein, WO die ID-Schwärme sinnvollerweise zusammen gesammelt werden.

Die wirklich interessante Frage finde ich, wie und durch wen die Berechnung des Taxon-Schwarms geschehen soll.

1. Die leichtestste Lösung wäre sicherlich eine Anfrage an die zentrale Datenbank, die ALLE relevanten Konzepte zurückgibt, aus denen sich der Client dann die für ihn relevanten raussucht. Ich denke, diesen Fall sollten wir auf jeden Fall implementieren, da er vom Client relativ wenig (Verständnis) verlangt. Diese Lösung würde auch Hierarchieinformationen in den Alternativklassifikationen miteinschliessen und wäre somit generell anwendbar.
2. Allerdings ist die Lösung natürlich mit einem Overhead verknüpft, da evtl. deutlich mehr Konzepte berechnet werden, als notwendig sind. Ein intelligenterer Client würde also im besten Fall die Konzeptbeziehungen für alternative Taxa gleich mitverwalten und könnte diese Berechnung dann selber durchführen. Allerdings mit dem von dir bereits skizzierten Problem, dass er auch alle Kindtaxa mit herunterladen und die Hierarchie aufbauen müsste. Ich gehe davon aus, dass die meisten Clients damit Probleme haben werden.
3. Ein mittel-intelligenter Client wüsste zumindest, von welchen Alternativklassifikationen er überhaupt Taxa führt und würde diese Information der Webservice-Anfrage an die zentrale Datenbank mitgeben, damit diese dann nur selektiv für die genannten Klassifikationen nach Konzept- und Kindrelationen sucht.

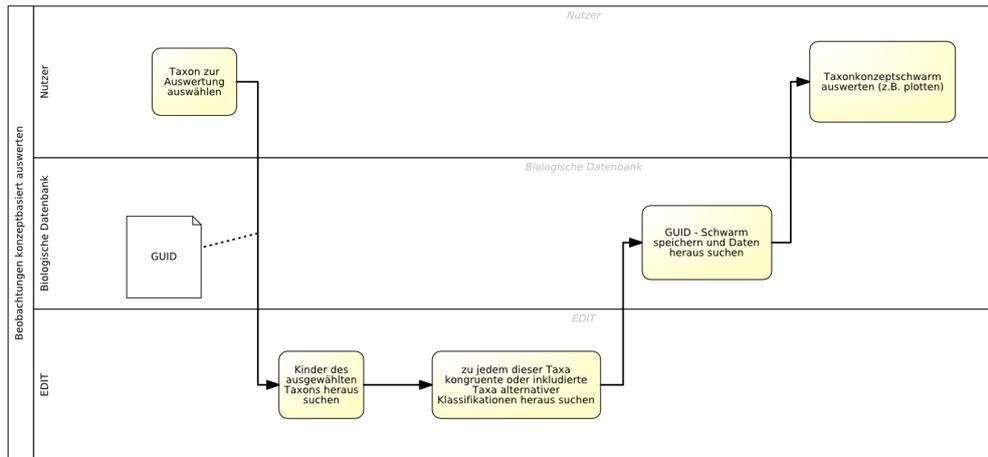


Abbildung 3: Vorgehen beim Auswerten von biologischen Datenbanken mit Hilfe konzeptverknüpfter Checklisten.

Für unser Beispiel bedeutet das. Ich möchte eine Verbreitungskarte von *Festuca ovina*, gehe also mit der ID 2 ins Rennen:

```
[1] 2
Festuca ovina has no childs.
ID-Swarm von Festuca ovina :
[1] 2 12
```

In unserem Beispielfall müssen für die Karte neben der ID 3 für *Festuca ovina L.* nach Wissk & Haeupler 1998, auch die ID 12, das ist *Festuca ovina subsp. ovina* nach Jäger & Werner 2005 heraus gesucht werden.

6.1 Nach Alternativkonzept auswerten

Wenn wir hingegen das Konzept von *Festuca ovina L.* aus Jäger & Werner 2005 verwenden wollen, z.B. weil wir bemerkt haben, dass dieses in unserer Funddatenbank viel häufiger verwendet wurde, so benötigen wir sowohl eine neue Konzeptrelationstabelle, als auch eine neue taxonomische Hierarchie und müssen damit die Schwarmbildung wiederholen.

TaxonUsageID	TaxonName	Classification	Relation	ID	AcceptedName
22	Festuca ovina	Ascherson 1864	=	11	Festuca ovina agg.
24	Festuca ovina subsp. duriuscula	Ascherson 1864	=	16	Festuca brevipila
23	Festuca ovina subsp. vulgaris	Ascherson 1864	>	12	Festuca ovina
23	Festuca ovina subsp. vulgaris	Ascherson 1864	<	11	Festuca ovina agg.
26	Festuca ovina var. tenuifolia	Ascherson 1864	=	15	Festuca filiformis
25	Festuca ovina var. vulgaris	Ascherson 1864	=	12	Festuca ovina
6	Festuca brevipila Wissk. u. Haeupler 1998		=	16	Festuca brevipila
5	Festuca filiformis Wissk. u. Haeupler 1998		=	15	Festuca filiformis
4	Festuca guestfalica Wissk. u. Haeupler 1998		=	14	Festuca ovina subsp. guestfalica
3	Festuca ovina Wissk. u. Haeupler 1998		=	13	Festuca ovina subsp. ovina
3	Festuca ovina Wissk. u. Haeupler 1998		<	12	Festuca ovina
2	Festuca ovina agg. Wissk. u. Haeupler 1998		=	11	Festuca ovina agg.

Childs of Festuca ovina (11):

TaxonUsageID	TaxonName	TaxonRank	AccordingTo	IsChildTaxonOfID	GENERATION	SYNONYM
3	12 Festuca ovina subsp. ovina	SSP	Jaeger & Werner (2005)	11	1	FALSE
4	13 Festuca ovina subsp. guestfalica	SSP	Jaeger u. Werner 2005	11	1	FALSE

Entsprechende oder inkludierte Konzepte:

TaxonUsageID	TaxonName	TaxonRank	AccordingTo
2	Festuca ovina	SPE	Wissk. u. Haeupler 1998
11	Festuca ovina	SPE	Jaeger u. Werner 2005
12	Festuca ovina subsp. ovina	SSP	Jaeger & Werner (2005)
13	Festuca ovina subsp. guestfalica	SSP	Jaeger u. Werner 2005
22	Festuca ovina	SPE	Ascherson 1864
23	Festuca ovina subsp. vulgaris	SSP	Ascherson 1864

zu verwendende Beobachtungen:

[1] 2 11 12 13 22 23

Zu diesem Konzept gibt es 6 Beobachtungen von 6 verschiedenen Taxonymen in unserer Beispieldatenbank (vergleiche auch mit Abbildung 1).

In den Angaben der Datenrepositories (z.B. GBIF, aber das gilt bisher genauso für Deutschlandflora.de) gehen die Konzepte vollkommen durcheinander (siehe ??). Die Angaben von *Festuca ovina* müssen in aller Regel zu *Festuca ovina* agg. gestellt werden. Um eine realistische Verbreitungskarte zu bekommen, müssten die Teilprojektdatenbanken (aus dem Mittelgebirge und Süddeutschland) ihre sicheren Angaben der Kleinart entsprechend deutlich ausweisen, und die Datencrawler müssten konzeptbasiert und nicht namensbasiert arbeiten.

7 Korrektur von Funddateninterpretationen und Fusionierung von Beobachtungsdatenbanken

In den bisherigen Überlegungen wurde ignoriert, dass es für die Mehrzahl der Primärdaten zuerst einmal unklar, zumindest aber undokumentiert ist, welchem Konzept die jeweiligen Namen folgen. In manchen (z.B. den auf den FLOREIN-ID's beruhenden, oder der GermanSL folgenden) ist dies zwar theoretisch klar, de facto beruhen aber fast alle Datenbanken auf Namens- und nicht auf Konzeptidentitäten.

Auf der anderen Seite fallen im Laufe des Zusammenspiels von Datenbanken zahlreiche Dinge auf, die man nicht auf der Ebene der Referenzlisten lösen möchte oder sollte. Zum Beispiel

1. Fehleingaben (zum Teil systematische),
2. seltene Misapplications, die auch nicht als auct.-Fälle in Referenzlisten integriert werden sollten

Es ist nicht sinnvoll, jede Misapplication zu einem auct. Fall in der Referenzliste zu machen. Wenn immer möglich, sollten diese Fälle in den Originaldatenbanken durch Reinterpretationen, z.B. mit Hilfe von Zuordnungstabellen korrigiert, im Falle von eindeutigen Fehleingaben auch hart verändert und dokumentiert werden.

An dieser Stelle auch noch einmal der deutliche Hinweis, dass nicht jeder Namensgebrauch der nicht mehr dem heutigen Standard entspricht, eine Misapplication ist. Meist ist lediglich unklar, welchem (älteren) Konzept diese Angabe entspricht. So ist in den Deutschlandatlaskommentaren oft von Verwechslungen die Rede, z.B. für *Festuca guestfalica*: „Die Karte bildet vielfach nur Kartieraktivitäten ab. In kalkreichen Gebieten sicher oft mit der sehr ähnlichen, aber azidophilen *Festuca ovina* s.str. verwechselt, d.h. eigentlich häufiger als dargestellt.“ In aller Regel dürfte sie in diesen „Verwechslungs“-Fällen aber nicht ausdrücklich als *Festuca ovina*-Kleinart kartiert worden sein, sondern ganz nackt als *Festuca ovina* L. In diesen Fällen muss man zuerst fragen, ob der Beobachter die Konzepte von *F. guestfalica* oder ihren Schwestertaxa *F. filiformis* und *F. brevipila* überhaupt kennt und (für seine Kartierung) akzeptiert. Ansonsten handelt es sich nämlich um eine korrekte Bestimmung eines akzeptierten Taxons (wenn auch aus einer veralteten Klassifikation).

7.1 Fusionierung von Daten unterschiedlicher Herkunft

Für Datensammelprojekte wie die Deutschlandfloraplattform (aber auch alle ernsthaft betriebenen Regionaldatenbanken) brauchen wir eine Zwischenschicht, in der alle Taxonnamen einem echten Konzept, aber nicht unbedingt der Standardsicht, zugeordnet werden. Fehlbestimmungen und -zuordnungen können im gleichen

Zuge, oder getrennt davon korrigiert werden. Beide Fälle müssen aber auf jeden Fall eindeutig dokumentiert werden.

Für Datensammelprojekte sollte eine Hierarchie von Korrekturtypen erarbeitet werden. Das heißt, nicht alle Anpassungen sollten auf Einzelfundebene behandelt werden, sondern wo immer möglich sollten sie taxonweise oder zumindest projektweise erfolgen, da dies nicht nur den direkten Korrekturprozess vereinfacht, sondern auch das Dokumentieren und Vermitteln der gewonnenen Einschätzungen ermöglicht.

8 Schlussbemerkungen

Das Zusammensammeln der ID-Schwärme kann entweder auf dem Server geschehen (hier stehen alle Konzept- und Hierarchieinformationen zur Verfügung) oder auf der Clientseite (hier steht, welche Alternativ-Konzepte relevant sind, weil sie in der Datenbank vorkommen). Wir sollten einfach ausprobieren, was praktikabler ist.

Zu klären ist, an welchen Stellen welche Informationen z.B. über die taxonomische Hierarchie notwendig sind, um dem Nutzer Anhaltspunkte über das Konzept eines Taxons jenseits des Klassifikationsnamens geben zu können.

Weitere Probleme der Erstellung von Verbreitungskarten, wie der Umgang mit regional monotypischen Taxa (*Poa trivialis* und *Poa trivialis* subsp. *trivialis* sind gültige Taxa, aber bezogen auf Deutschland inhaltlich identisch, da nur diese eine Unterart vorkommt), sind hier nicht erörtert.

Prinzipiell würde das hier skizzierte Vorgehen Synonyme für die Datenspeicherung weitestgehend verzichtbar machen, weil alle relevanten Synonyme besser als Teil einer zu integrierenden anderen Klassifikation abgelegt werden sollten. Aber das ist eher eine akademische als eine praxisrelevante Überlegung, so dass wir auch weiterhin Synonyme wie im Beispiel gezeigt in den Datenbanken nutzen und bei der Auswertung (!) zum Schwarm dazusammeln sollten.

Dies ist ein erster Entwurf für das Vorgehen, wie ich es mir vorstelle. Ich bin dankbar für alle Form von Rückmeldung, das allgemeine Vorgehen, genauso wie das vorliegende Dokument oder Details der Implementierung betreffend: f.jansen@bgbm.org