















































Siebenundzwanzigster Diversity Workbench Workshop am IT Zentrum der SNSB, 21.10.2015


I. ImportWizard anhand von Flora-von-Bayern-Datenbeispielen; DWB-konforme Standard-Excel-Tabellen (M. Ruff, M. Weiss) mit praktischen Übungen



Durchzuführende Schritte:

1. Kopfdaten der Excel Datei BFL_Workshop_Exceltab.xlsx ergänzen, so dass in jeder Zeile Kopfdaten enthalten sind. Die Dateien für diesen Workshop befinden sich im Ordner ...DiversityCollection/Import/Workshop2015.
2. Excel Tabelle als Text-Datei speichern (Text Tabstopp-getrennt). Dazu die ersten 5 Zeilen entfernen, so dass die Spaltenüberschriften in der ersten Zeile stehen.
3.  DiversityCollection öffnen
4. Mit Server verbinden: Auf  links oben klicken. **Name or IP-address of the Server:** 10.153.187.119; **Port:** 55 (siehe ausgeteilte Zettel); Login SQL-Server authentication anklicken; **User:** WORKSHOP, **Password:** dwb dann auf Connect to server klicken und bei Choose database DiversityCollection_Workshop Wählen.
5. Import Wizard öffnen: Data →  Import →  Wizard →  import Specimen
6. Text-Datei öffnen:  File →  um die Text Datei auszuwählen
7. Import Schema erstellen oder öffnen:
 - 7.1. Allgemeines:
 - Auf der linken Seite werden Felder gewählt, die importiert werden können. Auf der rechten Seite wird der Inhalt der Felder gewählt.
 - Über den Button  kann die Beschreibung des Feldinhalts für die jeweiligen Felder abgerufen werden
 - Es muss immer mindestens eine **decisive column** (?) gewählt werden. Decisive column bedeutet, dass immer dann Daten der Importdatei importiert werden, wenn in dieser Spalte Daten vorhanden sind. Sind in derselben Kategorie (z.B.  **Event** → **Date and time**) weitere Felder gewählt, ohne **decisive column** (?), werden diese nur mitimportiert, wenn das Feld mit der **decisive column** (?) einen Inhalt hat. Leere Felder werden grundsätzlich nicht importiert.
 - Rechts oben werden immer die für die jeweilige Kategorie noch fehlenden Angaben angezeigt.
 - 7.2. Zielprojekt auswählen:  **Project; ProjectID** → For all auswählen: Dort das jeweilige Projekt wählen, welchem die zu importierenden Daten zugeordnet werden sollen. In unserem Fall: **DWBworkshopBFL**. Diese Spalte als **decisive column** (?) auswählen
 - 7.3. Datum wählen: links bei  **Event** Haken setzen. Rechts **Event** → **Date and time**; für **Collection Day, Collection Month, Collection Year** Haken setzen. Jeweils From file wählen und die entsprechenden Datumsspalten aus der zu importierenden Datei zuordnen. Die **decisive column** (?) sollte für **Collection Month** und **Collection Year** gewählt werden.



- 7.4. Rasterfeld angeben: links  **MTB (A, CH, D)** wählen, rechts für **MTB (A, CH, D)**; **MTB** Haken setzen und  wählen, über From file die Spalte **TK** aus der Tabelle wählen und für **MTB (A, CH, D)**; **Quadrant** den **QQQ** aus der Tabelle wählen.
- 7.5. Kartierer verknüpfen: links  **Agent** wählen, rechts **Agent; CollectorsName**, From file: **Kartierer** wählen,  setzen.
- 7.6. Fundort: rechts  **Event; LocalityDescription**, From file: **Fundort** wählen,  setzen
- 7.7. Verwaltungseinheit: links  **Named area (Diversity Gazetteers)**, rechts **Named area (Diversity Gazetteers); Thes. code**, From file **ID Verwaltungseinheit** wählen,  setzen. Bei Pre.: <http://tnt.diversityworkbench.de/Gazetteer2/> einfügen. Dies ist der Ort wo die ID im Gazetteers gespeichert ist und somit der dahinterliegende Inhalt abgerufen werden kann.
- 7.8. Standort: rechts  **Event; HabitatDescription**, From file **Standort** wählen,  setzen.
- 7.9. Rechts- und Hochwert: links  **Gauss-Krüger coordinates**, rechts **Gauss-Krüger coordinates; R, H**, From file **Rechtswert** und **Hochwert** wählen,  setzen.
- 7.10. U [m] (Unschärfe): rechts  **Gauss-Krüger coordinates; LocationAccuracy**, From file **U [m]** wählen.
- 7.11. Sippenname nach Liste: links  **Organism** und  **Identification**, rechts **Organism** → **Identification; TaxonomicName**, From file **Sippenname nach Liste** wählen,  setzen. Bei **IdentificationCategory** For all: **determination** wählen. Zudem bei  **Organism; TaxonomicGroup**, For all: **Plant** wählen und bei **OnlyObserved** For all Haken setzen, wenn es sich um reine Beobachtungsdaten handelt.
- 7.12. Sippenname frei: wie 7.11, aber bei **TaxonomicName** auf  klicken und **Sippenname frei** aus Tabelle wählen.
- 7.13. LastIdentificationCache zuordnen: rechts  **Organism** → **Hierarchy, name; LastIdentificationCache** **Sippenname nach Liste** und  **Sippenname frei** anwählen,  setzen.
- 7.14. Status: links  **Analysis**, rechts **Analysis; AnalysisResult**, From file **Status** wählen,  setzen. Bei **AnalysisID**, For all, **Floristischer Status** wählen und bei **AnalysisNumber**, For all, **1** eintragen. Bei **Notes**, For all **Floristischer Status neu** eintragen. Falls zweiter floristischer Status vorhanden, links bei **Analysis** auf das  klicken und analog verfahren. Allerdings bei **AnalysisNumber** diesmal bei For all **2** eintragen.
- 7.15. Bestimmer: rechts  **Organism** →  **Identification** →  **Responsible; ResponsibleName**, From file **Bestimmer** wählen,  setzen.
- 7.16. Belegart: rechts  **Organism** → **Notes etc.; Notes**, From file **Belegart** wählen,  setzen. Zusätzlich bei Pre.: „Belegart: “ eintragen.
- 7.17. Belegnr.: Ebenfalls bei Notes. Hierzu das  wählen, für einen zweiten Notes Eintrag und bei Pre.: „; Belegnr.: “ reinschreiben.
- 7.18. Anzahl: rechts  **Organism; NumberOfUnits**, From file **Anzahl** wählen,  setzen.
- 7.19. Fläche:  **Organism** → **Notes etc.; Notes**. Zusätzlich bei Pre.: „; Fläche“ und bei Post: „m²“ reinschreiben.
- 7.20. Anmerkungen: → **Notes etc.; Notes**. Zusätzlich bei Pre.: „; Anmerkung“ reinschreiben.
- 7.21. Bisher wurde keine Auswahl für  **Specimen** getroffen. Hier könnte z.B. über **ExternalIdentifier** eine ID des Original Datensatzes importiert werden. Da wir diese nicht




haben, ein Import aber ohne Inhalt bei Specimen nicht funktioniert, wählt man für **ExternalIdentifier**, For all: **Test** aus und **?** setzen.







8. Mergen: Da Kopf- und Sammlungsdaten in einer Tabelle sind, die Struktur in der Datenbank aber so aufgebaut sein soll, dass alle Arten unter einem Sammlungsereignis (entspricht Kopfdaten) untergliedert werden sollen, muss die Datenbank die Information erhalten, welche Daten in ein Sammlungsereignis sollen. Dazu müssen die Kopfdaten abgeglichen werden. Bei gleichen Daten, werden die Einzelfunde den Kopfdaten untergliedert. Hierzu wird die Funktion  **Merge** in der DWB verwendet.









Unter  Merging für die Kopfdaten **Event**, **Gauss-Krüger coordinates**, **Named area (DiversityGazetteers)**, **MTB (A, CH, D)**, **Specimen**, **Agent 1**, **Project 1** Merge auswählen. Danach müssen die einzelnen Felder ausgewählt werden, welche abgeglichen werden sollen. Dies geschieht über das Symbol Key columns . Für folgende Felder müssen Key columns ausgewählt werden:

- **Event:** LocalityDescription und HabitatDescription
- **Event** → **Daten and time:** CollectionDay, CollectionMonth, CollectionYear
- **Gauss-Krüger coordinates:** R und H
- **Named area (Diversity Gazetteers):** Thes. code
- **MTB (A, CH, D):** MTB, Quadrant
- **Specimen** → **Notes:** InternalNotes. Dort die gesamten Kopfdaten aneinandergereiht auswählen, von Datum bis U [m]. Dies dient dazu, dass der Bereich Specimen eine einmalige Identifikation bekommt und dadurch abgeglichen werden kann.
- **Agent 1:** CollectorsName
- **Project:** ProjectID













9. Import-Schema testen: Um einzelne Zeilen zu testen wählt man  **Testing**. Bei **Test data in line** kann man die zu testende Zeile wählen. Danach durch klicken auf **Test data in line** ausführen. Nun wird einem für die gewählte Spalte angezeigt, ob es Probleme beim Import gibt und ob ggf. noch etwas vergessen wurde, wie z.B. Felder ohne **decisive column** .



















10. Fehlermeldungen bearbeiten: Es wurden nicht numerische Werte in das numerische Feld **NumberOfUnits** eingetragen. In diesem Fall ist eine Transformation der Daten notwendig. Hierzu auf den fehlerhaften Bereich gehen (rechts **Organism 1** → **NumberofUnits**). Dort auf  (Transformation) klicken und  (Replace parts of the value) wählen. Nun das zu ersetzenden Zeichen bei **Replace** eintragen und **With** leer lassen. Bei mehreren zu ersetzenden Zeichen muss mehrfach  gewählt werden. Unten auf **Test the transformation** klicken. Wenn alles stimmt, Fenster schließen.

11. Mehrere Einträge bei Kartierer trennen: Es kann immer nur ein Agent je Feld eingetragen werden. Um den Doppeleintrag zu trennen muss wie folgt vorgegangen werden. Bei **Agent 1** → **CollectorsName**,  wählen. Auf  (Split the value) klicken. Als **Splitter** „;“ nehmen und testen. Nun wird nur der erste Name angezeigt. Um den zweiten Teil anzeigen zu können muss man einen zweiten Agent anlegen und für diesen ebenfalls einen Splitter eintragen. Links auf  neben  **Agent 1** klicken. Rechts bei **Agent 2** → **CollectorsName** Kartierer auswählen und auf Transformation gehen. Bei **Position 2** wählen. Als **Splitter** wieder „;“ verwenden und testen. Um die Anführungszeichen zu entfernen entsprechend 10. vorgehen. Für Agent 2 muss das Merging noch ergänzt werden. Sowohl unter  Merging bei Agent 2 Merge anwählen, als auch unter Agent 2 den  anklicken.

12. Schema speichern: unter  **Import** → auf **Save import schema** klicken.
13. Test-Import: Den ersten Import möglichst nicht im Zielprojekt testen, sondern z.B. auf einem Workshop Projekt. Es könnten Fehler bei dem dann nötigen Löschen des Imports entstehen. Zum Import auf  **Import** →  **Start import** klicken. Bei jedem Import wird ein Import Schema und ein Import Protokoll gespeichert.
14. Falls nach dem Test-Import noch Fehlermeldungen angegeben werden, diese beheben und danach erneut importieren.
15. Schema öffnen und Datensatz importieren: Hat man einen Datensatz, der von der Struktur zu einem bestehenden Schema passt, ist es nicht notwendig eine neues Schema zu erstellen. Hierzu zunächst die Text-Datei öffnen:  **File** →  um die Text Datei auszuwählen. Dann bei Schema auf  klicken um ein Schema zu öffnen. Wie unter 9. Testen und anschließend importieren. Falls die Daten schon geprüft wurden auf  **Import** →  **Start import** klicken.

II. DiversityCollection-Suchabfragen zur Qualitätskontrolle und Datenanalyse; Datenexporte (M. Ruff, M. Weiss) mit praktischen Übungen

1. Sippen-Namen mit Taxref abgleichen und verbinden:
 - 1.1. Dazu muss zunächst eine Verbindung mit dem Modul DiversityTaxonNames hergestellt werden. **Connection** →  **Module Connection** → **Diversity TaxonNames** auf  klicken (Falls DiversityTaxonNames Plants bereits vorhanden und rosa hinterlegt, auf  klicken und das Passwort eingeben). **Name:** tnt.diversityworkbench **Port:** 5432, **User:** tnt, **Password:** mu7idSwg, **Database:** DiversityTaxonNames_Plants
 - 1.2. Im nächsten Schritt werden die importierten Namen mit den Namen der Taxon Referenzliste abgeglichen. **Administration** → **Database** → **Maintenance**. Dann **Synchronize Database** → **Collection** <-> **TaxonNames** → **Synchronize taxonomic names missing a connection:** Bei **Taxonomy database** DiversityTaxonNames_Plants on tnt.diversityworkbench.de wählen. Bei **Taxonomy project:** BFLnames wählen, bei **Project** DWBWorkshopBFL wählen, bei **TaxonomicGroup** plant wählen. Zunächst compare whole name wählen. Hier werden die Sippennamen nach Liste gefunden. Die Sippennamen frei weichen in der Regel zumindest geringfügig von den Namen in TaxonNames ab. Deshalb vergleicht man nur die ersten Teile der Namen. Hierzu Compare first parts wählen und mit 4, 3, 2 und zuletzt 1 versuchen. Update similar name wählen (dadurch wird der alte Name überschrieben, wird dies nicht gewünscht Insert as new name wählen). Dann auf Check for differences klicken. In der Ergebnisliste wird angezeigt, wie viel Entsprechungen es jeweils in der Taxonomischen Referenzliste gibt. Diese können mit  **Select valid name** angezeigt und ausgewählt werden.
2. Für die Namen der Kartierer analog vorgehen über **Collection** <-> **Agents**, zudem mit TNT Server für Agents verbinden. Bei Table einmal für CollectionAgent (für den Kartierer) und einmal Identification (für den Bestimmer) durchführen. Bei Display type alle relevanten Namensschreibweisen durchführen.
3. Für die Landkreis ID **Collection** <-> **Gazetteers**, Synchronize places with a link to the Diversity Gazetteers verwenden, zudem mit TNT Server für Gazetteers verbinden.
4. Es bleiben in der Regel einige Namen übrig, die nicht zugeordnet werden konnten. Diese findet man, in dem man nach leeren Links sucht. Dies muss für **TaxonomicNames, Collectors** und **Named area** durchgeführt werden. Über das Symbol  auf der Hauptseite können Suchfelder ergänzt werden. Es werden die Felder **Identification** → **Name URI und Collector** → **Name URI** benötigt. Im Suchkriterien Feld jeweils \emptyset wählen, bei **ordnen nach:** wählen, was angezeigt werden soll. Für die Suche nach den nicht verlinkten Artnamen z.B. **Tax. name** wählen. Nun können durch klicken auf  die nicht verbundenen Daten gesucht werden. Im Anschluss können diese direkt in DiversityCollection abgeändert werden. Die Suche nach fehlenden Links zum Gazetteer muss noch programmiert werden.
5. Für die Quadranten sollten die Koordinaten der vier Eckpunkte generiert werden. Diese sind nötig um die Quadranten in der Karte darstellen zu können. **Maintenance** → **Collection events** → **Coordinates for TK25**. Source: DiversityGazetteers2, Project: DWBWorskhopBFL.
6. Kartenausgabe zur Kontrolle: In einem Datensatz das Feld MTB anklicken. Oben rechts die  Kartendarstellung aktivieren. Statt  Browser  GIS – Edit auswählen und statt  Collection event  Distribution wählen. Bei den beiden sich öffnenden Fragefenstern jeweils auf Ja klicken. Danach auf  und die Daten der Karte hinzufügen.

7. Suchabfragen zur Datenanalyse:
 - 7.1. Datum: z.B. - auswählen für Daten zwischen zwei Tagen, > für alle Daten nach einem Bestimmten Datum.
 - 7.2. Statistik über Projekte aufrufen: **Help** →  **Statistics** → **Project**: Projekt auswählen, dann z.B.  Collectors und  Taxonomy aufrufen. Projektübersichten können über **Projects** angezeigt werden.
 - 7.3. Suchabfragen mit anschließender Kartendarstellung: z.B. einzelne Arten, einzelne Projekte, bestimmte Zeiträume. Nach der Suchabfrage für die Kartenerstellung wie unter 6. Vorgehen.
8. Export-Wizard
 - 8.1. Vor dem Export muss sichergestellt werden, dass eine Verbindung mit den benötigten Modulen hergestellt wurde. Siehe 1.1.
 - 8.2. Die zu exportierenden Daten müssen über die Suchfunktion ausgewählt werden. Für ein ganzes Projekt z.B. das entsprechende Projekt auswählen und auf  klicken.
 - 8.3. Um in den Export-Wizard zu gelangen **Data** →  **Export** →  **Wizard** →  **Organisms** wählen
 - 8.4. Man sollte sich zunächst überlegen, welche Daten man in welcher Reihenfolge exportieren möchte. Um einen besseren Überblick über die zu exportierenden Daten und deren Feldzuweisung zu erhalten, ist es hilfreich sich ein zweites Fenster mit DiversityCollection und demselben Projekt zu öffnen.
 - 8.5. Auf der linken Seite des Export-Wizards können die Überkategorien der Daten gewählt werden. Auf der rechten Seite wählt man die zu exportierenden Felder. Um ein Feld hinzuzufügen wählt man auf der rechten Seite das  bei dem jeweiligen Feld.
 - 8.6. Bei Feldern die mit einem externen Modul verlinkt sind wie z.B. Collector oder Tax. name können beliebige zusätzliche Informationen aus dem verknüpften Modul ausgewählt werden. Hierzu wählt man unterhalb dem zu exportierenden Feld , danach wählt man die Datenbank mit welcher man die Inhalte verlinkt hat und zuletzt, was aus dieser Datenbank exportiert werden soll. Um die Daten dem Export hinzuzufügen muss man noch auf das  neben dem eben gewählten klicken.
 - 8.7. In der Regel beginnt man mit den Kopfdaten. Für das Datum würde man beispielsweise links **Event** wählen und rechts bei **CollectionDay, CollectionMonth, CollectionYear** jeweils auf das  klicken. So geht man ähnlich dem Import, bloß in umgekehrter Reihenfolge, Stück für Stück vor.
 - 8.8. Will man beim Export die aktuellste Nomenklatur der Taxa verwenden, muss man links **Identification** wählen und rechts unter **NameURI**  DiversityTaxonNames_Plants wählen. Dort Accepted name wählen und auf das  klicken. Dadurch werden nur aktuell gültige Namen ausgegeben.
 - 8.9. In der Zeile unten bei  **Test** lässt sich zum überprüfen immer eine aktuelle Vorschau des Exports anzeigen.
 - 8.10. Schema speichern: unter  Schema den Speicherpfad wählen und auf  Speichern klicken.
 - 8.11. Export: Auf  Export gehen, Verzeichnis und Dateinamen wählen und auf Export Data  klicken. Es wird eine Text Datei generiert. Der Inhalt dieser Datei kann nach Excel kopiert werden.