

|                        |    |       |                 |
|------------------------|----|-------|-----------------|
| Entomologica Austriaca | 14 | 17-42 | Linz, 23.2.2007 |
|------------------------|----|-------|-----------------|

**Zur Algorithmierbarkeit von Synonymieproblemen in  
Biodiversitätsdatenbanken.  
Anforderungsanalyse und praktische Lösungsvorschläge am  
Beispiel der Chrysomeliden (Insecta: Coleoptera)**

E. GEISER & M. MALICKY

**Abstract:** The Feasibility of Algorithmic Solutions for Problems with Synonyms in Biodiversity Databases. An Analysis of Requirements and Practical Solutions exemplified by the Chrysomelids (Insecta: Coleoptera). Synonyms – using different scientific names for one and the same species – generates a lot of problems when storing and retrieving species-specific data in biodiversity databases. We have developed a technique for an automatic or semi-automatic correlation of the scientific names of any one species used which are synonymous with the currently valid scientific name. The technique can be carried out with minimum time and effort for storing and retrieving species-specific information. It was implemented successfully in the ZOBODAT database.

First, requirements were analysed to draw up the criteria needed for an algorithmic solution to the synonym problem. The most relevant criteria are:

- It must be possible to store specimens using both the current valid scientific name or any synonyms formerly used.
- Changes in the future of the scientific names have to be implemented in an easy way.
- There must be no loss of information whatsoever when analysing all synonyms of a species.
- The databank should be both easy to program and maintain.
- Any non-specialist must be able to store data easily.

Exemplified by the Chrysomelids of Central Europe, their synonym problems were divided up into 8 separate classes. For categories 1 to 7 a practical solution was presented and tested to see if it met requirements. The synonyms in category 8 have to be handled by experts.

Possible solution for categories 1 to 5:

Each scientific name is given a number, its characteristic number. Each scientific name is then given a second synonym number. If a scientific name is currently valid, the characteristic number and the synonym number are identical. If a scientific name is a synonym, the value of its synonym number is the characteristic number of the valid name of the same species.

Possible solution for categories 6 and 7:

These categories comprise the problems which arise then species are split up, while the name of the original species does not change. The new separated species gets a new scientific name and can be labelled with a new characteristic number. But all records stored with the characteristic number of the original species now refer to two species. This characteristic number will now be flagged with "agg." meaning "aggregate". To characterise specimens with an old name, but new sense (without specimens of the separated species), the former unchanged scientific name is given a new characteristic number, which is flagged with "sensu stricto". In the database table the close relation of these three (or more!) different numbers has to be mapped.

The way in which these numbers (in the different synonym categories) are allocated for the various categories is exemplified by the Chrysomelids species of Central Europe. The frequency distribution of the synonym problems related to the species number shows that with 84 % of the species synonym problems could be easily dealt with using the above algorithm and an additional 10 % of the species could be partially dealt with. Other than the Chrysomelids, most of the insect species in Central Europe show the same frequency distribution of their synonym problems, particularly the Coleoptera and Lepidoptera species, which make up a high percentage of the records on invertebrate data in biodiversity databases. Therefore, our technique is very useful both for storing and retrieving with minimal time and effort and without loss of information. An additional advantage is that the storage and retrieval can be mostly performed by non-specialists. Our technique provides a step by step guide including which cases require specialists or where specialist supervision is necessary.

**Key words:** systematics, biodiversity, synonymy, storage and retrieval in databases, algorithms, requirement analysis.

## Inhalt

|      |                                                                                                                                                                    |    |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1.   | Einleitung.....                                                                                                                                                    | 19 |
| 1.1. | Begriffserläuterung.....                                                                                                                                           | 19 |
| 2.   | Sinn und Zweck von Biodiversitätsdatenbanken .....                                                                                                                 | 20 |
| 3.   | Grundlegendes Tabellendesign .....                                                                                                                                 | 21 |
| 4.   | Das Hauptproblem: der dynamische Kenntniszuwachs in der Systematik versus dauerhafte Speicherung von artbezogenen Informationen aus verschiedenen Jahrzehnten..... | 22 |
| 5.   | Die Chrysomeliden Mitteleuropas als Modellgruppe für Synonymieprobleme in Datenbanken .....                                                                        | 23 |
| 6.   | Algorithmierbarkeit .....                                                                                                                                          | 24 |
| 6.1. | Automatisierung.....                                                                                                                                               | 24 |
| 6.2. | Anforderungsanalyse .....                                                                                                                                          | 25 |
| 7.   | Kategorisierung der verschiedenen Problemfälle.....                                                                                                                | 26 |
|      | Kategorie 1: Verwendung mehrerer Gattungs- und/oder Artnamen bei ein und derselben Art.....                                                                        | 26 |
|      | Kategorie 2: Unterschiedliche Schreibweisen.....                                                                                                                   | 29 |
|      | Kategorie 3: Aus Unterarten werden Arten.....                                                                                                                      | 30 |
|      | Kategorie 4: Von einer bisher monotypischen Art wird eine neue Unterart erstellt.....                                                                              | 30 |

|                                                                                                   |    |
|---------------------------------------------------------------------------------------------------|----|
| Kategorie 5: Bisher als getrennt betrachtete Arten werden zu einer einzigen zusammengezogen ..... | 31 |
| Kategorie 6: Arten werden aufgespalten .....                                                      | 32 |
| Kategorie 7: Aus Arten werden Unterarten .....                                                    | 36 |
| Kategorie 8: Automatisierung nicht sinnvoll .....                                                 | 36 |
| 8. Häufigkeitsanalyse der verschiedenen Fälle .....                                               | 37 |
| 9. Schlussfolgerungen .....                                                                       | 39 |
| 10. Danksagung .....                                                                              | 40 |
| 11. Zusammenfassung .....                                                                         | 40 |
| 12. Literatur .....                                                                               | 41 |

## 1. Einleitung

Synonyme, verschiedene wissenschaftliche Namen für ein und dieselbe Art, gehören zum Alltag eines Systematikers, aber auch zum Alltag aller Biologen, deren wissenschaftliche Tätigkeit sich auf konkrete Arten bezieht. Während die Zahl der untersuchten Arten bei Physiologen oder Verhaltensforschern meist relativ gering und damit auch in ihrer Systematik überschaubar bleibt, hat man als Ökologe oder Tiergeograph mit einer derartigen Vielzahl von Arten zu tun, dass eine ständige Aktualisierung der verwendeten Namen unverhältnismäßig viel Forschungszeit und -geld beanspruchen würde. Zudem wird der Informationsgehalt der Ergebnisse dadurch meist in keiner Weise gesteigert! Von der Synonymieflut besonders betroffen sind die Biodiversitätsdatenbanken.

Während Systematiker im allgemeinen durchaus einsehen, dass zoologische Sammlungen oder die Herbarbelege in den Museen nicht alle 5 bis 10 Jahre neu etikettiert und umgeordnet werden können, wird ein solcher Anspruch durchaus für Biodiversitätsdatenbanken erhoben, in der irrigen Annahme, dass man das mit dem Computer "mit ein paar Mausklicks" erledigen könne. Dabei werden zwei wesentliche Dinge übersehen:

- Wenn der Anwender (neudeutsch: User) bei einem Computerprogramm ganz einfach einige Tasten drückt oder ein paar Mal mit der Maus klickt, dann verbergen sich dahinter sehr komplexe Programmabläufe, die zuvor aufwändig erstellt wurden.
- Ein Problem ist nur dann mit einem Computer "per Mausklick" lösbar, wenn die Aufgabenstellung algorithmierbar ist. Das bedeutet, dass sich der Lösungsweg vollständig in einzelne logische und eindeutige Anweisungen zerlegen lässt. Das ist bei den Synonymieproblemen nur teilweise der Fall. Welche Synonymieprobleme und in welcher Form sie algorithmierbar sind, wird im Folgenden am Beispiel der Chrysomeliden Mitteleuropas untersucht und diskutiert.

### 1.1. Begriffserläuterung

**Binärname und Artname:** Üblicherweise werden mit dem Begriff "Artname" sowohl beide Wörter der binären Nomenklatur (z. B. "*Leptinotarsa decemlineata*" für den Kartoffelkäfer) als auch nur das zweite Wort ("*decemlineata*") bezeichnet. Die

genaue Bedeutung ergibt sich aus dem Zusammenhang. Im vorliegenden Artikel müssen wir aber zwischen den beiden Bedeutungen auch in einem einzigen Satz genau unterscheiden. Zur besseren Verständlichkeit nennen wir daher eine wissenschaftliche Artbezeichnung aus Gattungs- und Artnamen "Binärnamen". Mit "Artnamen" meinen wir also immer nur das zweite Wort des Binärnamens. Zur Kennzeichnung einer Art genügt also der "Artnamen" alleine nicht!

**Spezialist:** Mit "Spezialist" bezeichnen wir im Folgenden einen sehr gut eingearbeiteten Kenner einer überschaubaren Organismengruppe, der sowohl die Arten dieser Gruppe zuverlässig determinieren kann, aber vor allem auch die aktuelle systematische Situation dieser Gruppe genau kennt. Ein "Spezialist" kann sowohl männlich als auch weiblich sein. Es kommt nicht auf das Geschlecht, sondern nur auf die Fachkenntnisse an!

**Algorithmierbarkeit:** Im Folgenden bezeichnen wir ein Synonymieproblem als algorithmierbar, wenn die Zuordnung der unterschiedlichen Binärnamen ein und derselben Art zueinander vollständig durch Datenbankdesign und -programmierung eindeutig und korrekt durchgeführt werden kann.

## 2. Sinn und Zweck von Biodiversitätsdatenbanken

Biodiversitätsdatenbanken enthalten typischerweise folgende Information:

Welche konkrete Art wurde an welchem konkreten Ort an welchem Datum festgestellt (beobachtet oder gesammelt) plus Angaben zur Zuverlässigkeit und Überprüfbarkeit dieses Sachverhaltes: Wer war der Sammler bzw. der Beobachter, wer hat das festgestellte Individuum determiniert, wo befindet sich der Beleg bzw. wo wurde diese Meldung publiziert.

Die Einrichtung und der Betrieb von Biodiversitätsdatenbanken sind aufwändig und teuer. Dennoch werden sie in zunehmendem Maße erstellt und betrieben, weil man diese artbezogenen Daten für verschiedene Zwecke auswerten möchte.

### Die wesentlichen Auswertungsziele sind dabei

- Erstellung von Artenlisten eines geographischen Gebietes.
- Erstellung von Verbreitungsinformationen über einzelne Arten, wenn möglich auch als Verbreitungskarten (Punkt- oder Rasterkarten).
- Komplexe Auswertungen, z. B. Phänogramme, Berechnung von Schutzwürdigkeitsindizes und anderen Kennziffern, GIS-Auswertungen.
- Bereitstellung von Informationen für systematische Revisionen, also die Antwort auf die Frage: Wo befinden sich die üblichen Verdächtigen, die ich als Spezialist noch genauer untersuchen muss?
- Naturschutzfachliche Auswertungen: Naturschutzinteressen stehen oft im Gegensatz zu ökonomischen Interessen. Naturschutzargumente benötigen daher eine solide, nachvollziehbare Datenbasis. Die Möglichkeit der gezielten Erhebung von Verbreitungsdaten von Arten der FFH-Richtlinie oder der diversen regionalen Roten Listen und deren Auswertung ist eines der Hauptmotive von

Naturschutzbehörden, Biodiversitätsdatenbanken zu betreiben bzw. finanziell zu unterstützen.

- Monitoring von potentiellen Krankheitsvektoren, Neozoa und Neophyta: Dieser Bereich gewinnt zunehmend an Bedeutung. Von einigen der jüngst eingewanderten bzw. eingeschleppten Arten sind negative Auswirkungen bekannt, daher ist deren Ausbreitungsverhalten von besonderem Interesse. Andere stehen unter Beobachtung, weil man noch nicht sicher weiß, ob sie negative Auswirkungen haben werden. Ob eine – meist vorschnell postulierte – Korrelation von Ausbreitungstendenzen und Klimawandel besteht, kann bei diesen und anderen Arten nur nach entsprechenden Auswertungen aus Biodiversitätsdatenbanken nachgewiesen werden.

Alle diese Auswertungen liefern aber nur dann sinnvoll interpretierbare Ergebnisse, wenn die zugrundeliegende Datenbasis repräsentativ, also einigermaßen umfangreich ist! Als Grundprinzip gilt: Je mehr Daten – und zwar aus möglichst verschiedenen Zeiträumen – gespeichert sind, umso mehr Auswertungen sind möglich und umso zuverlässiger sind die Ergebnisse. Beispiele für Auswertungsmöglichkeiten und die wesentlichen Kriterien für den erfolgreichen Betrieb einer Biodiversitätsdatenbank wurden in GEISER (1998, 2005) publiziert. Um die oben definierten Auswertungsziele zu erreichen, ist die dauerhafte Speicherung von artbezogenen Informationen aus vielfältigen (aber zuverlässigen) Datenquellen aus einem längeren Zeitraum (30 bis über 100 Jahre) notwendig.

### 3. Grundlegendes Tabellendesign

Biodiversitätsdatenbanken werden zweckmäßigerweise mit einem relationalen Datenbankmanagementsystem (im Folgenden vereinfacht "relationale Datenbank" genannt) eingerichtet. Eine relationale Datenbank enthält mehrere Tabellen, die so gestaltet sind, dass die Informationen immer nur an jeweils einer einzigen Stelle gespeichert werden, aber durch Verknüpfung der Information in den Tabellenfeldern sehr komplexe Speicher- und Auswertungsmöglichkeiten geschaffen werden. Wie man die meist sehr komplexen Aufgabenstellungen aus der realen Welt möglichst zweckmäßig in Tabellenfeldern zusammenfasst und diese Felder dann möglichst geschickt auf die einzelnen Tabellen verteilt, ist eine Kunst für sich, die nur wenige Informatiker beherrschen!

Eine Biodiversitätsdatenbank enthält – neben anderen für ihren Betrieb notwendigen und nützlichen Tabellen – immer die folgenden, hier vereinfacht dargestellten Tabellen:

#### Arten-Tabelle

| Binärnamen-Nr | Gattungsname       | Gattung_Autor | Artname                | Art_Autor_Jahr |
|---------------|--------------------|---------------|------------------------|----------------|
| 65506         | <i>Chrysomela</i>  | LINNÉ         | <i>vigintipunctata</i> | (SCOPOLI 1763) |
| 65508         | <i>Chrysolina</i>  | MOTSCHULSKY   | <i>fastuosa</i>        | (SCOPOLI 1763) |
| 65485         | <i>Crepidodera</i> | CHEVROLAT     | <i>aurata</i>          | (MARSHAM 1802) |

**Orte-Tabelle**

| Orte-Nr | Fundort-Bezeichnung | Ort     | (Bundes)Land | geogr. Länge | geogr. Breite | Höhe üNN |
|---------|---------------------|---------|--------------|--------------|---------------|----------|
| 5083    | Gartenau            | Grödig  | Salzburg     | 13° 2,5'     | 47° 44,4'     | 458 m    |
| 5084    | St. Leonhard        | Grödig  | Salzburg     | 13° 2,6'     | 47° 44,6'     | 456 m    |
| L125    | St. Leonhard        | Tamsweg | Salzburg     | 13° 48,2'    | 47° 7,3'      | 1110 m   |

Die Arten-Tabelle und Orte-Tabelle gehören zu den Grunddaten-Tabellen. In diesen Tabellen muss es eine Spalte geben, in der die Einträge alle voneinander verschieden sind. An Hand dieses einen Eintrags kann man dann eine ganze Tabellenzeile eindeutig identifizieren. Diese besondere Spalte nennt man Primärschlüssel einer Tabelle. In diesen beiden Beispielen fungiert jeweils die erste Spalte als Primärschlüssel. Bei der Speicherung einer konkreten Fundmeldung genügt es dann, bei der Art die zugehörige Binärnamen-Nummer und die Ort-Nummer des Fundortes in die Fundmeldungen-Tabelle einzutragen:

**Fundmeldungen-Tabelle**

| Binärnamen-Nr | Orte-Nummer | Fund-Datum | Sammler   | Determinator | Beleg        |
|---------------|-------------|------------|-----------|--------------|--------------|
| 65506         | 5083        | 23.05.2003 | E. Geiser | E. Geiser    | coll. Geiser |
| 65508         | 5083        | 12.07.2002 | E. Geiser | E. Geiser    | coll. Geiser |
| 65485         | 5083        | 23.05.2001 | E. Geiser | E. Geiser    | coll. Geiser |

Anmerkung: Die Felder "Sammler", "Determinator" und "Beleg" werden in der Praxis ebenfalls durch Grunddatentabellen mit Primärschlüsseln implementiert.

Bei einer Auswertung wird dann über die Binärnamen-Nummer der Binärname eruiert und über die Orte-Nummer die genauen Ortsdaten.

**Auswertungsbeispiel**

Käferfunde in der Gemeinde Grödig, nach dem Jahr 2002 und unter 1000 m ü.N.N.

*Chrysomela vigintipunctata*, 23.05.2003, Beleg in coll. Geiser

*Chrysolina fastuosa*, 12.07.2002, Beleg in coll. Geiser

#### **4. Das Hauptproblem: dynamischer Kenntniszuwachs in der Systematik versus dauerhafte Speicherung von artbezogenen Informationen aus verschiedenen Jahrzehnten**

Eine Biodiversitätsdatenbank, mit der man sinnvolle Auswertungen erstellen kann, enthält artbezogene Informationen aus vielen Jahrzehnten. Nun haben sich aber im Lauf der Jahre die Namen vieler Arten geändert, bei manchen Arten sogar mehrmals. Über Sinn und Unsinn dieser Namensänderungen kann man in jedem einzelnen Fall lange diskutieren. Aber die Synonyme sind eine Tatsache, für die ein Datenbank-Betreiber eine praktikable Lösung entwickeln muss.

Um artbezogene Informationen in Biodiversitätsdatenbanken abspeichern zu können, müssen die Arten eindeutig gekennzeichnet sein<sup>1</sup>. Für die eindeutige Kennzeichnung weist man den verschiedenen in der Datenbank zu speichernden Einheiten eindeutige Nummern zu, den sogenannten Primärschlüssel. Nun wäre es natürlich naheliegend, jeder Art eine solche eindeutige Nummer zuzuweisen. Im Unterschied zu einem Systematiker weiß ein Computer aber nicht, was eine Art ist. Man kann daher nur jedem Binärnamen einen Primärschlüssel zuweisen. Hat eine Art mehrere Binärnamen, so vergibt man für jeden dieser Namen eine eigene Nummer. Damit ist es bereits möglich, artbezogene Informationen unter den verschiedenen Binärnamen, die für diese Art verwendet werden, korrekt und ohne Informationsverlust einzuspeichern.

Sobald man aber versucht, artbezogene Information auszuwerten (was ja der eigentliche Zweck einer solchen Datenbank ist), hat man ein Problem. Es ist daher unerlässlich, die Information mit abzuspeichern, welche Binärnamen Synonyme von ein und derselben Art sind und welcher dieser Namen als aktuell gültig angesehen wird. Dieses Problem potenziert sich noch, sobald man versucht, Biodiversitätsdatenbanken zu vernetzen und gemeinsame Auswertungen oder den Druck von Verbreitungskarten aus der verteilten Datenmenge zu realisieren, obwohl das sehr sinnvoll und wünschenswert wäre. Während die einzelnen Datenbankbetreiber (hoffentlich) einigermaßen einen Überblick über die von ihnen verwendeten Systeme haben, treten beim Vernetzen von mehr als zwei Datenbanken sehr rasch massive Inkompatibilitäten auf. Das ist eine wesentliche, wenn auch nicht die einzige Ursache dafür, dass die bisherigen Vernetzungsversuche in den Anfangsstadien geblieben sind und sich auch nur auf ganz wenige, systematisch unproblematische Arten beschränkt haben

(z. B. EIS.net: [www.eis-international.org/index.php](http://www.eis-international.org/index.php) und GBIF: [www.gbif.org](http://www.gbif.org)).

## **5. Die Chrysomeliden Mitteleuropas als Modellgruppe für systematische Probleme in Datenbanken**

1987 erschien der Katalogband von LUCHT, eine Zusammenstellung der Binärnamen und Verbreitungsangaben nach den Bänden 1-11 des Standardwerkes "Die Käfer Mitteleuropas" (FHL 1-11). Bis 1994 erschienen drei Ergänzungsbände (FHL 12-14) und 1998 der Ergänzungsband dieser Ergänzungsbände (FHL 15). Von den rund 600 Namen mitteleuropäischer Chrysomeliden aus dem Lucht-Katalog blieb knapp die Hälfte von 1987 bis 1998 stabil. Seit 1998 sind weitere Namensänderungen erfolgt und vor 1987 wurden ebenfalls bei manchen Arten andere Namen verwendet. Das bedeutet, dass bei mehr als der Hälfte der mitteleuropäischen Chrysomelidenarten in den letzten 50 Jahren mehr als ein Binärname in Gebrauch war oder ist.

Diese Situation in Hinblick auf die Synonyme ist bei den Chrysomeliden keineswegs extrem, sondern typisch für viele mitteleuropäische Insektengruppen, insbesondere für Käfer und Schmetterlinge. Gerade von diesen beiden Insektenordnungen sind aber besonders viele artbezogene Daten vorhanden und werden auch in Biodiversitätsdatenbanken eingespeichert. Daher ist es unbedingt notwendig, eine adäquate Lösung der Syno-

---

<sup>1</sup> Das war auch vor über 250 Jahren die Intention eines Carl Linnaeus, der mit seiner binären Nomenklatur die Arten eindeutig benennen wollte. Er kann nichts dafür, dass man inzwischen seine Erfindung zur Verschlüsselung phylogenetischer Information zweckentfremdet.

nymieprobleme in einer Biodiversitätsdatenbank zu implementieren.

Die Autoren sind durch ihre jahrzehntelange Arbeit an der Datenbank ZOBODAT ([www.zobodat.at](http://www.zobodat.at), MALICKY 2001) und bei der Erstellung einer Käferfaunistik des Landes Salzburg (GEISER 2001) und der Checkliste der Chrysomeliden Österreichs (GEISER 2004) massiv mit den hier geschilderten Problemen konfrontiert worden. Vergleicht man den Aufbau und die Verfahrensweise anderer Biodiversitätsdatenbanken und diskutiert man mit den Betreibern, so muss man feststellen, dass eine adäquate Lösung für das Problem der Synonyme noch nicht gefunden bzw. umgesetzt worden ist.

Nun ist Synonym aber nicht gleich Synonym, wie jeder Systematiker weiß. Es gibt einfache, eindeutige Namensänderungen und solche, die daraus resultieren, dass Arten aufgespalten, Teile von Arten umgruppiert oder versteckte Arten entdeckt wurden. Um die Synonyme in einer Biodiversitätsdatenbank sinnvoll zu handhaben, haben die Autoren die Probleme in verschiedene Kategorien unterteilt und anschließend für jede Kategorie eine möglichst praktikable Lösung erarbeitet und in der Datenbank ZOBODAT realisiert. In manchen dieser Fälle kann man die Synonymisierung zur Gänze an den Computer delegieren, in anderen ist es immerhin partiell möglich. In einigen Fällen muss die genaue Zuordnung aber ausschließlich durch Spezialisten durchgeführt werden.

## 6. Algorithmierbarkeit

### 6.1. Automatisierung

Der hohe Lebensstandard in der westlichen Zivilisation beruht zu einem erheblichen Teil auf der Tatsache, dass in den letzten Jahrhunderten und vor allem Jahrzehnten immer mehr Routinetätigkeiten weitgehend automatisiert wurden, von der Mühlentechnik über die industrielle Fertigung bis zum Kauf inklusive Abrechnung von Flugtickets und Opernkarten.

Bevor man einen Prozess automatisiert, muss man vorher folgendes abklären:

- a. Welche Ziele möchte man durch Automatisierung erreichen?
- b. Aus welchen Teilprozessen besteht der ganze Prozess?
- c. Welche Teilprozesse sind automatisierbar?
- d. Welche Teilprozesse sind partiell automatisierbar?
- e. Welchen Aufwand verursacht die Automatisierung im Verhältnis zu den definierten Zielen?

Punkt a bezeichnet man auch Anforderungsanalyse. Für die Punkte c und d muss anschließend ein Algorithmus (logische Folge von eindeutigen Anweisungen) gefunden werden, der zu einem erwünschten Ergebnis führt. Den Punkt e nennt man Kosten-Nutzen-Analyse. Sehr oft stellt sich dabei heraus, dass man mit moderatem Aufwand (= Kosten) bereits 80 % oder 85 % der erwünschten Ziele erreichen kann, während die Erfüllung weiterer Ziele nur mit einem Vielfachen des bisher kalkulierten Aufwands erreichbar ist. Weiters steigt dadurch die Komplexität des Verfahrens deutlich an, was zusätzlich einen erhöhten Wartungsaufwand, längere Schulungsmaßnahmen und größere Fehleranfälligkeit des Gesamtsystems zur Folge hat.

## 6.2. Anforderungsanalyse

Eine Lösung für die Synonymieprobleme bei Biodiversitätsdatenbanken muss folgende Anforderungen erfüllen:

1. Möglichkeit der Einspeicherung von aktuellen Binärnamen aus der Literatur und aktuell revidierten Sammlungen.
2. Möglichkeit der Einspeicherung von Binärnamen aus der "älteren" Literatur ("älter" bedeutet: älter als 2 Jahre!).  
Je älter, umso mehr weicht die Nomenklatur von der aktuellen ab. Andererseits wurden in den letzten 10 bis 100 Jahren zahlreiche hervorragende Faunen- und Florenwerke erstellt, die wertvolle Informationen beinhalten.
3. Möglichkeit der Einspeicherung aus gut bestimmten Sammlungen.  
Das wesentliche Qualitätskriterium einer Sammlung ist nicht, ob sie der neuesten Systematik bzw. der in der Datenbank verwendeten Systematik entspricht, sondern ob sie zum Zeitpunkt ihrer Aufstellung mit einem dem damaligen Kenntnisstand entsprechenden hohen Zuverlässigkeitsgrad determiniert worden ist. Dazu gehören vor allem die Privatsammlungen von Spezialisten, die oft eine eigene Systematik verwenden, die sich mit der in der Datenbank üblichen nicht immer einfach zur Deckung bringen lässt. Gerade diese Sammlungen sind aber besonders wertvoll (ausführlich in GEISER 1996).
4. Die Namensänderungen müssen nachvollziehbar sein.
5. Es darf kein Informationsverlust auftreten!!
6. Der weitere Fortschritt in der Systematik muss implementierbar sein!  
Auch in Zukunft ist mit weiteren Synonymisierungen zu rechnen. Bei vielen Gruppen ist der derzeitige systematische Kenntnisstand unbefriedigend. Außerdem werden durch die modernen molekularbiologischen Methoden ständig neue Systematik-Kenntnisse gewonnen, versteckte Arten entdeckt oder Umgruppierungen nahegelegt. An dieser Stelle soll aber auch daran erinnert werden, dass viele Namensgebungen keineswegs auf harten, belegbaren Fakten beruhen, sondern Auffassungssache sind. Es gibt zahlreiche alte – und leider auch weniger alte – Artbeschreibungen, die einen Interpretationsspielraum offen lassen. Der zu einer Abklärung erforderliche Holotypus ist oft nicht mehr vorhanden. Bei den Gattungsnamen verschärft sich noch das Problem: Für die Zusammenfassung von Arten zu einer Gattung gibt es zwar viele gute klassische und molekulare Kriterien, letztlich ist die Zuordnung aber eine Entscheidung des Bearbeiters, die jederzeit von einem anderen Bearbeiter abgeändert werden kann.
7. Es muss eine artbezogene Auswertung durch alle Synonymieketten möglich sein. Viele wesentliche Auswertungen beziehen sich auf Bestandsveränderungen in den letzten Jahrzehnten, also über einen Zeitraum, in dem für viele Arten mehrere Synonyme in Gebrauch waren. Die Auswertungen sollen dennoch korrekte Ergebnisse liefern.
8. Das Datenbankdesign einer Lösung muss so gestaltet sein, das der Programmieraufwand für die Datenspeicherung und Auswertung und der Wartungsaufwand minimal sind.
9. Für die Vernetzung mit anderen Datenbanken müssen genau definierte Schnittstellen bereitgestellt werden.

10. Der Einspeicherungsmodus muss praktikabel sein und von Nichtspezialisten durchgeführt werden können!

Dieser letzte Punkt ist eigentlich der wichtigste. Der Einspeicherungsmodus ist ein kritischer Faktor, der über Erfolg oder Misserfolg einer Biodiversitätsdatenbank entscheidet! Wenn man für die Einspeicherung eines artbezogenen Datensatzes statt einer Minute 5 oder 10 Minuten braucht, so bedeutet das, dass mit demselben (Zeit- oder Geld-)Aufwand nur ein Zehntel der Daten gespeichert werden können! Andererseits sind bei einer zu geringen Datenbasis viele Auswertungen wertlos. Die Bedeutung des Einspeicherungsmodus und anderer kritischer Erfolgsfaktoren einer Biodiversitätsdatenbank wurden in GEISER (2005) ausführlich erläutert.

Wesentlich ist ebenfalls, dass die Dateneingabe überwiegend von Nichtspezialisten durchgeführt werden kann, allerdings unter der Aufsicht bzw. in Zusammenarbeit mit Spezialisten der jeweiligen Gruppe. Die Zeit der Spezialisten ist angesichts der Fülle ihrer Aufgaben immer beschränkt, ob sie einem ganz anderen Brotberuf nachgehen oder einen Teil, aber eben nur einen Teil(!) ihrer Dienstzeit darauf verwenden dürfen. Sie darf daher keineswegs für die Routine-Dateneingabe missbraucht werden. Es bleiben genug problematische Fälle, die nur der Spezialist bearbeiten kann!

## 7. Kategorisierung der verschiedenen Problemfälle

Synonymieprobleme treten in Biodiversitätsdatenbanken dann auf, sobald für eine Art mehr als ein Binärname verwendet wird. Die Gründe dafür sind unterschiedlich. Im Folgenden werden die einzelnen Synonymie-Ursachen untersucht und eine Lösung für deren Handhabung in einer Datenbank vorgestellt.

### Kategorie 1

#### Verwendung mehrerer Gattungs- und/oder Artnamen bei ein und derselben Art

##### Fallbeispiele für Kategorie 1

##### Synonymie der Artnamen

*Chaetocnema picipes* (STEPHENS 1831)  
 = *Chaetocnema laevicollis* (THOMSON 1866)  
 = *Chaetocnema heikertingeri* LJUBISCHEFF 1963

Diese Art wurde mehrfach beschrieben. Bei einer späteren Revision wurde erkannt (eigentlich: postuliert), dass diese Namen alle dieselbe Art bezeichnen. Laut Nomenklaturregeln wurde daher das älteste Synonym zum (derzeit!) gültigen Namen erklärt.

##### Synonyme der Gattungsnamen

*Macrolea appendiculata* (PANZER 1794)  
 = *Haemonia appendiculata* (PANZER 1794)

Dieser Fall ist bei den Chrysomeliden und bei den meisten anderen Organismengruppen

sehr häufig. Manchmal werden aus formaljuridischen Gründen sämtliche Arten einer Gattung mit einem anderen Gattungsnamen belegt. Ein weiterer Grund sind die Auffassungsunterschiede der einzelnen Bearbeiter, welche Arten in einer Gattung zusammengefasst werden sollen. Die üppige Fülle an Gattungsnamen, insbesondere bei einigen umfangreichen Chrysomelidengattungen, findet zwar nicht immer Eingang in die Standardwerke. Sobald neue Namen aber publiziert sind, sind sie existent und werden – in unterschiedlichem Ausmaß – eine Zeit lang verwendet.

Verwirrender wird es noch, wenn die Gattungsnamen nicht nur geändert, sondern auf eine andere Gattung übertragen werden:

*Chrysolina fastuosa* (SCOPOLI 1763)

= *Dlochrysa fastuosa* (SCOPOLI 1763)

= *Chrysomela fastuosa* (SCOPOLI 1763)

= *Fastuolina fastuosa* (SCOPOLI 1763)

*Chrysolina herbacea* (DUFTSCHMID 1825)

= *Chrysomela herbacea* DUFTSCHMID 1825

= *Chrysomela menthastri* SUFFRIAN 1851

*Chrysomela vigintipunctata* (SCOPOLI 1763)

= *Melasoma vigintipunctata* (SCOPOLI 1763)

*Linnaeidea aenea* (L. 1758)

= *Melasoma aenea* (L. 1758)

= *Chrysomela aenea* L. 1758

Diese Binärnamen sind zwar eindeutig einer Art zuordenbar, aber diese kleine Beispielsreihe zeigt bereits, wie akribisch man bei der Datenspeicherung vorgehen muss. Dass diese diffizilen Namensänderungen einige der häufigsten mitteleuropäischen Chrysomelidenarten betreffen, macht die Sache nicht besser.

Die bisher geschilderten Fälle haben allerdings den großen Vorteil, dass eine umkehrbar eindeutige Namenszuordnung zur selben Art möglich ist. Durch diese umkehrbar eindeutige Zuordnungsmöglichkeit wird dieser Synonymiefall algorithmierbar.

### Konkrete Lösung bei der Speicherung in der Datenbank

Jeder Binärname erhält eine eigene Binärnamen-Nummer in der Datenbank, die gleichzeitig als Primärschlüssel der Arten-Tabelle fungiert. Zusätzlich muss die Information in der Datenbank abgespeichert werden, welcher der derzeit gültige Name ist und welche Namen zu diesem synonym sind. Das erfolgt durch die Vergabe einer Synonymie-Nummer in einer weiteren Tabellenspalte, die nach folgendem Schema aus der Binärnamen-Nummer erstellt werden: Beim (derzeit) gültigen Namen wird die Binärnamen-Nummer als Synonymie-Nummer vergeben, wodurch der gültige Name in der Datenbank auf einfache Weise gekennzeichnet wird. Bei den Synonymen verwende man als Synonymie-Nummer die Binärnamen-Nummer des gültigen Namens.

**Arten-Tabelle**

| Binärnamen-Nr | Synonym-Nr | Gattungsname      | Artname           |
|---------------|------------|-------------------|-------------------|
| 65397         | 65397      | <i>Chrysolina</i> | <i>herbacea</i>   |
| 39789         | 65397      | <i>Chrysomela</i> | <i>herbacea</i>   |
| 65548         | 65397      | <i>Chrysomela</i> | <i>menthastri</i> |

Die Auswertung lässt sich dann sehr einfach programmieren.

Pseudocode:

//drucke gültigen Binärnamen aus:

gültiger Binärname : wenn Binärnamen-Nr = Synonym-Nr

//werte alle Datensätze dieser Art aus

suche alle Datensätze mit derjenigen Binärnamen-Nr, die dieselbe  
Synonymie-Nr (wie vorhin gespeichert) haben

**Nun muss noch geprüft werden, ob das hier geschilderte Verfahren die Anforderungskriterien erfüllt:**

ad 1: Datenerfassung unter dem aktuellen Namen.

Das ist problemlos möglich.

ad 2 und 3: Datenerfassung aus älterer Literatur und aus verschiedenen alten Sammlungen.

Da jedes Synonym eine eigene Binärnamen-Nummer erhält, ist das problemlos möglich. Jede Art wird einfach unter dem Namen eingespeichert, mit dem sie in der Sammlung steckt oder publiziert wurde. Bei der Datenerfassung ist es irrelevant, ob der angegebene Binärname gültig oder ein Synonym ist. Die Einspeicherung artbezogener Information ist so unter jedem Synonym möglich. Ein heute gültiger Binärname kann schließlich schon morgen ein Synonym sein!

ad 4 und 5: Namensänderungen müssen nachvollziehbar sein, es darf kein Informationsverlust auftreten.

Die Fundmeldungen werden immer unter der Binärnamen-Nummer gespeichert, die dem Binärnamen der Datenquelle (Sammlung oder Publikation) entspricht. Dadurch bleibt die Information über die ursprüngliche Bezeichnung erhalten. Durch die Synonymienummer ist die Namensänderung nachvollziehbar.

ad 6: Der weitere Fortschritt in der Systematik muss implementierbar sein.

Falls sich bei einer späteren Revision herausstellt, dass ein anderer, älterer oder ganz neuer Binärname Gültigkeit haben sollte, müssen nur die Synonym-Nummern in der Artentabelle ausgetauscht werden. Man muss also nur eine einige Nummern an einer einzigen Stelle in der Datenbank ändern! Alle artbezogenen Informationen, die unter der Binärnamen-Nummer in den anderen Tabellen abgespeichert sind, bleiben unverändert. Ab nun sind Auswertungen mit dem neuen gültigen Namen und mit weiterhin korrekter Synonymiezuordnung möglich!

ad 7: Artbezogene Auswertung muss über alle Synonyme möglich sein.

Durch die gemeinsame Synonymie-Nummer kann man bei jeder Auswertung sämtliche

Synonyme mit einer einfachen Datenbankabfrage problem- und verlustlos dem jeweils gültigen Namen zuordnen.

ad 8: Minimaler Programmieraufwand.

Bei dem hier beschriebenen Verfahren kann man sämtliche Auswertungsprogramme unverändert weiterhin verwenden, auch wenn neue Synonyme hinzugefügt werden oder ein anderer Name die aktuelle Gültigkeit erhält.

ad 9: Bereitstellung von genau definierten Schnittstellen für die Vernetzung mit anderen Datenbanken.

Bei einer Vernetzung mit anderen Datenbanken muss man davon ausgehen, dass unterschiedliche Namen für jeweils eine Art verwendet werden. Da mit dem hier beschriebenen Verfahren alle Synonyme einer Art als zusammengehörig gekennzeichnet sind, ist die Vernetzung über ein beliebiges Synonym möglich. Falls die andere Datenbank einen Namen verwendet, der sich noch nicht in der Synonymliste befindet, erweitert man die eigene Datenbank einfach um diesen weiteren Namen (siehe Kriterium 6).

ad 10: Datenerfassung durch Nichtspezialisten.

Der Datenerfasser muss nicht wissen, ob der abzuspeichernde Binärname ein Synonym ist. Wenn der Binärname in der Liste der speicherbaren Namen aufscheint, dann kann er verwendet werden!

## Kategorie 2

### Unterschiedliche Schreibweisen

Durch Schreibfehler, Druckfehler oder Änderung bzw. striktere Auslegung der Nomenklaturregeln fanden unterschiedliche Schreibweisen für denselben Binärnamen Eingang in die Literatur und auch in jahrzehntelang verwendete Bestimmungswerke. Bei den Revisionen der letzten Jahrzehnte wurden einige dieser Schreibfehler aufgedeckt und korrigiert, sie sind aber auf Determinationsetiketten und in der Fachliteratur weit verbreitet. Eine einfache Korrektur in der Artentabelle der Datenbank wäre nicht zweckmäßig. Das widerspricht den Kriterien 4, 7, 8, 9 und 10.

### Fallbeispiele für Kategorie 2

Es ist für Nichtspezialisten nicht einfach zu beurteilen, ob eine abweichende Schreibweise einer ähnlich geschriebenen Art zuzuordnen oder doch eine andere Art ist:

*Altica* = *Haltica*, aber ungleich *Halictus*

Die ersten beiden Namen bezeichnen dieselbe Gattung der Erdflöhe (Ordnung: Coleoptera), *Halictus* bezeichnet eine Gattung der Schmalbienen (Ordnung: Hymenoptera).

*Cryptocephalus pallifrons* = *Cryptocephalus pallidifrons*

aber:

*Cryptocephalus nitidus* ungleich *Cryptocephalus nitidulus*

Die Ähnlichkeiten der Bezeichnung wie

*Donacia malinovskyi* und *Donacia malinowskyi*

oder

*Mantura mathewsii* und *Mantura matthewsi*

oder

*Pachybrachys hippophaes* und *Pachybrachys hippophaes*

legen zwar den Verdacht der Bezeichnung für dieselbe Art sehr nahe. Dennoch bleibt hier vor allem das Problem bei der Auswertung bestehen: Sobald die Namen unterschiedlich geschrieben werden, erhält man kein oder ein verfälschtes Ergebnis.

Für die Handhabung in einer Biodiversitätsdatenbank ist es daher am zweckmäßigsten, solche unterschiedlichen Schreibweisen desselben Binärnamens wie einfache Synonyme der Kategorie 1 zu behandeln. Dadurch wird auch die Unsicherheit und Fehlerhäufigkeit bei der Dateneingabe deutlich reduziert. Weiters ist eine problemlose Auswertung unabhängig von der Schreibweise möglich.

| Binärnamen-Nr | Synonym-Nr | Gattungsname    | Artname          |
|---------------|------------|-----------------|------------------|
| 28095         | 28095      | <i>Epitrix</i>  | <i>atropae</i>   |
| 22398         | 28095      | <i>Epithrix</i> | <i>atropae</i>   |
| 28096         | 28096      | <i>Aphthona</i> | <i>czwalinae</i> |
| 5049          | 28096      | <i>Aphthona</i> | <i>czwalinai</i> |

Sämtliche Anforderungskriterien werden analog wie bei Kategorie 1 erfüllt.

### Kategorie 3

#### Aus Unterarten werden Arten

Manche aktuelle Arten wurden jahrzehntelang als gut unterscheidbare Unterarten angesehen. Wenn bisher in der Datenbank immer nur diese eine Unterart (trinäre Nomenklatur) und nie die Nominatsubspezies (weil letztere z. B. in Mitteleuropa nicht vorkommt) oder keine nur bis zur Art determinierten Exemplare (mit binärer Nomenklatur) gespeichert wurden, kann die Methode der Synonymnummern, die bei Kategorie 1 beschrieben wurde, auch hier angewandt werden.

#### Fallbeispiele für Kategorie 3

| Binärnamen-Nr | Synonym-Nr | Gattungsname       | Artname          | Unterartname     |
|---------------|------------|--------------------|------------------|------------------|
| 24008         | 24008      | <i>Longitarsus</i> | <i>cerinthes</i> |                  |
| 5085          | 24008      | <i>Longitarsus</i> | <i>nervosus</i>  | <i>cerinthes</i> |

Auch in diesen Fällen sind alle Anforderungen erfüllt, insbesondere die problemlose Einspeicherung älterer Datensätze, die noch als Subspezies determiniert wurden. Ebenso ist die Auswertung nach der aktuellen Systematik problemlos möglich.

### Kategorie 4

#### Von einer bisher monotypischen Art wird eine neue Unterart erstellt.

Das eben ausgeführte Verfahren funktioniert auch eine Ebene tiefer, bei der Erstellung einer neuen Unterart aus einer forma. Die neue Subspezies erhält eine neue Binärnamen-

Nummer (auch wenn es sich hier genau genommen um einen trinären Namen handelt). Weiters müssen die Grunddaten der neuen Nominatspezies verfügbar sein. Die Nummer des Binärnamens der Art braucht nicht geändert zu werden.

#### Fallbeispiel für Kategorie 4:

*Phratora polaris leederi* STEINHAUSEN, 1993

galt früher als Forma:

*Phyllodecta polaris* (SCHNEIDER, 1886) forma *leederi*

| Binärnamen-Nr | Synonym-Nr | Gattungsname       | Artname        | Unterartname   |
|---------------|------------|--------------------|----------------|----------------|
| 52216         | 52216      | <i>Phratora</i>    | <i>polaris</i> |                |
| 4958          | 52216      | <i>Phyllodecta</i> | <i>polaris</i> |                |
| 53097         | 53097      | <i>Phratora</i>    | <i>polaris</i> | <i>leederi</i> |
| 53098         | 53098      | <i>Phratora</i>    | <i>polaris</i> | <i>polaris</i> |

Da unter *Phratora polaris* bereits bisher immer die ganze Art inklusive aller früheren und zukünftigen Unterarten verstanden wurde, bleibt deren Binärnamen-Nummer unverändert. Durch den inzwischen geänderten Gattungsnamen ist hier eine Synonymisierung nach Kategorie 1 notwendig. Die neuen Unterarten erhalten neue Nummern.

#### Kategorie 5

##### Bisher als getrennt betrachtete Arten werden zu einer einzigen zusammengezogen

Die intraspezifische Variationsbreite ist nicht immer leicht abzuschätzen. Auch wenn der Erstbeschreiber bereits eine Serie der neuen Art untersucht hat, können in den darauf folgenden Jahren Übergangsformen zu bereits bekannten Arten entdeckt werden. So kann ein späterer Bearbeiter zu dem Schluss kommen, dass diese bisher als verschiedene Arten angesehenen Exemplare doch alle zu einer einzigen Art gehören. (Noch später kann sich allerdings herausstellen, dass die ursprüngliche Separierung der einen Art berechtigt war. Das kommt zwar nicht sehr häufig, aber doch immer wieder vor.)

##### Fallbeispiele für Kategorie 5

*Plateumaris discolor* PANZER 1795 wird nun als zur Art *Plateumaris sericea* (L., 1758) gehörig angesehen. *Plateumaris discolor* wird dadurch zum Synonym. Die "Vereinigung" der beiden Arten ist in der Datenbank durch Synonymie-Nummern problemlos möglich. Als *Plateumaris discolor* ausgewiesene Sammlungsexemplare oder Literaturmeldungen können immer noch unter dem ursprünglichen Binärnamen abgespeichert werden. Dadurch tritt kein Informationsverlust auf.

| Binärnamen-Nr | Synonym-Nr | Gattungsname       | Artname           |
|---------------|------------|--------------------|-------------------|
| 6578          | 6578       | <i>Plateumaris</i> | <i>sericea</i>    |
| 23406         | 6578       | <i>Plateumaris</i> | <i>discolor</i>   |
| 6579          | 6579       | <i>Plateumaris</i> | <i>rustica</i>    |
| 4771          | 6579       | <i>Plateumaris</i> | <i>affinis</i>    |
| 4927          | 4927       | <i>Phaedon</i>     | <i>armoraciae</i> |
| 6532          | 4927       | <i>Phaedon</i>     | <i>veronicae</i>  |

Ebenso wird *Plateumaris affinis* (KUNZE 1818) inzwischen als zu *Plateumaris rustica* (KUNZE, 1818) gehörig abgesehen. Ein weiteres Beispiel ist *Phaedon veronicae* BEDEL 1892, der seit Jahrzehnten als artgleich zu *Phaedon armoraciae* (L., 1758) angesehen wird, was allerdings FUSS et al. 2005 anzweifeln. Hier ist es besonders wichtig, dass alle als *Phaedon veronicae* determinierten Exemplare weiterhin unter eben diesem Namen abgespeichert werden.

Wie gut das Verfahren mit den Synonymie-Nummern funktioniert, kann hier noch am Beispiel der folgenden drei Namen demonstriert werden: Die Art *Lamprosoma kolbei* wird als artgleich zu *Lamprosoma concolor* angesehen. Der Gattungsname selbst ist synonym zu *Oomorphus*, der derzeit als gültig betrachtet wird. Bei *Neocrepidodera* liegt ein analoger Fall vor.

| Binärnamen-Nr | Synonym-Nr | Gattungsname          | Artname            |
|---------------|------------|-----------------------|--------------------|
| 65514         | 65514      | <i>Oomorphus</i>      | <i>concolor</i>    |
| 4862          | 65514      | <i>Lamprosoma</i>     | <i>concolor</i>    |
| 4863          | 65514      | <i>Lamprosoma</i>     | <i>kolbei</i>      |
| 65628         | 65628      | <i>Neocrepidodera</i> | <i>melanostoma</i> |
| 53231         | 65628      | <i>Asioestia</i>      | <i>melanostoma</i> |
| 5121          | 65628      | <i>Crepidodera</i>    | <i>melanostoma</i> |
| 5119          | 65628      | <i>Crepidodera</i>    | <i>frigida</i>     |

Auch hier ist mit der Methode der Synonymienummern eine eindeutige und informationsverlustfreie Abspeicherung unter dem jeweiligen Artnamen möglich und die Auswertung auf dem aktuellen Stand der Namenszuordnung durchführbar.

Falls diese zusammengezogenen Arten später wieder aufgespalten werden, können die unter den getrennten Nummern gespeicherten Datensätze sofort wieder verwendet werden. In den Datensätzen der unter der "Gesamtart" gespeicherten Meldungen sind allerdings beide Arten enthalten. Dieser Fall führt uns zur nächsten Kategorie.

## Kategorie 6

### Arten werden aufgespalten

Durch neue Erkenntnisse in der Systematik kommt es immer wieder vor, dass sich manche Arten als Komplex mehrerer biologischer Arten herausstellen und daher aufgespalten werden. In manchen Fällen hat man die morphologisch von der Nominatform abwei-

chenden Individuen bereits einer "forma" zugeordnet. Aber manchmal sind die morphologischen Unterschiede der neuen Art sehr subtil. Artunterschiede können sich speziell bei den Chrysomeliden auch in unterschiedlichen Futterpflanzen manifestieren (GROSS et al. 2004).

Die abgespaltene Art erhält üblicherweise einen neuen Artnamen und dann kann für sie sofort eine neue Nummer vergeben werden. Der Binärname der Art, von der diese neue Art abgespalten wurde, ändert sich allerdings nicht. Das Dilemma für die Datenbank besteht nun darin, dass – oft zahlreiche – Datensätze mit Artnamen gespeichert sind, die sich eigentlich auf zwei (oder mehr) Arten beziehen können. Weiters muss es ab nun möglich sein, neu determinierte Exemplare der "alten" Art, die sicher nicht zur abgespaltenen Art gehören, eindeutig einzuspeichern. Das darf nicht mehr unter der alten Binärnamen-Nummer geschehen, da diese ab nun die "vermischten" Datensätze kennzeichnet.

Bei der Abspaltung einer Art von einer anderen sind nun 3 verschiedene Binärnamen-Nummern als Primärschlüssel notwendig:

X1 für die abgespaltene Art

X2 für Exemplare der Art mit dem bisher verwendeten Binärnamen, die aber sicher nicht zum abgespaltenen Taxon gehören: in der Tabelle werden diese Arten im Artnamen als *sensu stricto* (s. str.) gekennzeichnet.

X3 Die alte Nummer des ursprünglichen Binärnamens wird weiterverwendet, der Artname wird in der Arten-Tabelle durch "agg." (Aggregat) ergänzt. Damit wird gekennzeichnet, dass unter dieser Nummer auch Exemplare gespeichert sind, die zur abgespaltenen Art gehören können.

Als Synonymie-Nummer wird die jeweilige Binärnamen-Nummer verwendet, sofern es sich bei allen 3 Binärnamen um gültige Namen handelt. Mit diesen Binärnamen-Nummern ist es möglich, sämtliche Exemplare korrekt einzuspeichern, die neue, abgespaltene Art und revidierte Exemplare der alten Art, die sicher nicht zur abgespaltenen Art gehören. Alle Fundmeldungen vor der Artrevision und solche, bei denen man nicht weiß, ob sie nach dem neuen Schlüssel determiniert wurden, werden mit der Aggregat-Nummer gespeichert.

Bei diesem Fall verwendet man drei Nummern für zwei Arten. Allerdings steckt in diesen Nummern nirgends die Information, dass und wie sie zusammengehören. Das ist bei der Einspeicherung noch kein Problem, aber für die Auswertung benötigt man noch eine weitere Tabelle, in der genau diese Information enthalten ist.

#### Aggregat-Tabelle

| Binärnamen-Nr | Aggregat-Nr |
|---------------|-------------|
| X1            | X3          |
| X2            | X3          |
| X3            | X3          |

In dieser Tabelle wird die Information gespeichert, dass die Arten mit den Nummern X1 und X2 aus der Art mit der Nummer X3 hervorgegangen sind.

**Fallbeispiele für Kategorie 6***Zeugophora frontalis* SUFFRIAN 1840wurde früher als eine forma von *Zeugophora scutellaris* SUFFRIAN 1840 angesehen.*Phaedon salicinus* HEER 1845früher: *Phaedon armoraciae* (L. 1758) variatio *salicinus*

| Binärnamen-Nr | Synonym-Nr | Gattungsname      | Artname                    |
|---------------|------------|-------------------|----------------------------|
| 62433         | 62433      | <i>Zeugophora</i> | <i>frontalis</i>           |
| 58321         | 58321      | <i>Zeugophora</i> | <i>scutellaris</i> s. str. |
| 4774          | 4774       | <i>Zeugophora</i> | <i>scutellaris</i> agg.    |
| 62440         | 62440      | <i>Phaedon</i>    | <i>salicinus</i>           |
| 31667         | 31667      | <i>Phaedon</i>    | <i>armoraciae</i> s.str.   |
| 4927          | 4927       | <i>Phaedon</i>    | <i>armoraciae</i> agg.     |

Einträge in der Aggregat-Tabelle

| Binärnamen-Nr | Aggregat-Nr |
|---------------|-------------|
| 62433         | 4774        |
| 58321         | 4774        |
| 4774          | 4774        |
| 62440         | 4927        |
| 31667         | 4927        |
| 4927          | 4927        |

Dieses Verfahren erfüllt zwar nicht mehr alle, aber doch sehr viele der Anforderungen.

ad 1. Möglichkeit der Einspeicherung von aktuellen Binärnamen.

Das ist durch die Verwendung der Nummern X1 und X2 problemlos möglich.

ad 2. u. 3. Möglichkeit der Einspeicherung von Binärnamen aus der "älteren" Literatur und aus gut bestimmten Sammlungen.

Immer, wenn man nicht sicher weiß, ob bei dieser Art agg. oder s.str. gemeint ist, speichert man die Meldung mit Aggregat-Nummer ab.

ad 4. Die Namensänderung ist nachvollziehbar.

Das wird durch die Aggregat-Tabelle gewährleistet.

ad 5. Es darf kein Informationsverlust auftreten.

Die vor der Revision erstellten Datensätze wurden mit der Aggregat-Nummer gespeichert. Jetzt weiß man zwar nicht, welche dieser Datensätze zu Exemplaren der abgespaltenen Art oder zur s.str. Art gehören, aber das wusste man bei der Determination dieser Exemplare auch nicht. Daher tritt auch kein Informations-

verlust auf. Allerdings ist ab jetzt ein neu entstandener Informationsbedarf vorhanden.

- ad 6. Der weitere Fortschritte in der Systematik muss implementierbar sein.

Gerade Artaufspaltungen werden durch die neuen Methoden und die weiteren Fortschritte in der Systematik noch oft vorkommen. Mit der vorgestellten Methode kann man sie in einer Biodiversitätsdatenbank handhaben.

- ad 7. Es muss eine artbezogene Auswertung über alle Synonymieketten hindurch möglich sein.

Diese Anforderung wird nur teilweise erfüllt! Da für denselben Binärnamen zwei verschiedene Nummern als gültig gekennzeichnet sind, ist bei der Auswertung Vorsicht geboten. Je nach Auswertungsziel gilt es abzuwägen, welche der beiden Nummern man verwenden möchte und in welcher Kombination.

- ad 8. Minimierung des Programmieraufwandes:

Diese Anforderung ist erfüllt. Wir kennen kein Verfahren, das programmiertechnisch noch einfacher ist.

- ad 9. Bereitstellung von genau definierten Schnittstellen für die Vernetzung mit anderen Datenbanken.

Man kann bei den Schnittstellen sowohl den Artnamen im aktuellen Sinn als auch im früheren Sinn anbieten.

- ad 10. Der Einspeicherungsmodus muss praktikabel sein und von Nichtspezialisten durchgeführt werden können.

Jeder Binärname hat seine Nummer und im Zweifelsfall nimmt man die Aggregat-Nummer. Allerdings darf das nicht dazu führen, dass eindeutig als s. str. determinierte Exemplare dann unter der Aggregat-Nummer abgespeichert werden. Gerade bei dieser Problemkategorie wird sehr deutlich, wie wichtig eine Supervision der Dateneingabe durch Spezialisten ist.

Bei dieser Kategorie Artaufspaltung wird deutlich, dass man die Synonymisierung nicht mehr vollständig an den Computer delegieren kann. In der Datenbank kann man zwar die vermischten Datensätze in einer Aggregat-Nummer zusammenfassen, aber diese Information muss dann von einem Spezialisten mit "Brainware" weiterbearbeitet werden. Der Spezialist erhält somit eine Übersicht, wo sich die verdächtigen Exemplare befinden, und kann sie revidieren – theoretisch. In der Praxis werden die Exemplare der abgespeicherten Datensätze nicht mehr vollständig vorhanden sein. Außerdem ist der Revisionsaufwand enorm und wird daher nur punktuell und nur bei manchen Fragestellungen stattfinden. Somit werden in den Datenbanken viele Datensätze in der Kategorie agg. verbleiben, Tendenz steigend!

Wichtiger als eine vermeintlich saubere Lösung (also die Revision sämtlicher Datensätze mit Aggregat-Nummern) ist es, die Neueinspeicherung aktuell determinierter Exemplare zu ermöglichen, ohne dass deren Namenszuordnung im alten Binärnamen der Sammelart untergeht!! Weiters muss man bei den Auswertungen bedenken, inwiefern man im konkreten Fall das Problem der "Sammelarten" berücksichtigen muss.

**Kategorie 7****Aus Arten werden Unterarten****Fallbeispiel für Kategorie 7**

Auch der folgende Fall muss mit dem Aggregat-Verfahren gehandhabt werden: Die neue Unterart *Clytra quadripunctata appendicina* erhält eine eigene Nummer, mit identischer Synonymie-Nummer, wodurch sie als aktuell gültig gekennzeichnet wird. Der frühere Binärname *Clytra appendicina* behält seine alte Binärnamen-Nummer, aber er erhält als Synonymienummer die der neu erstellten Unterart. Weiters muss man die Artentabelle mit den Daten der Nominatspezies inklusiver neuer Nummer ergänzen. Die alte Binärnummer von *Clytra quadripunctata*, die Daten ohne *Clytra appendicina* kennzeichnet, wird mit der Binärnummer der Nominatspezies synonymisiert. Da man aber nicht weiß, ob in Zukunft auch alle Exemplare dieser Art bis auf die Unterart determiniert werden, muss man eine neue Nummer für *Clytra quadripunctata* agg. einrichten, damit diese Daten nicht mit der alten, exakteren *Clytra quadripunctata* vermischt werden.

| Binärnamen-Nr | Synonym-Nr | Gattungsname  | Artname               | Unterartname          |
|---------------|------------|---------------|-----------------------|-----------------------|
| 65469         | 65469      | <i>Clytra</i> | <i>quadripunctata</i> | agg.                  |
| 4793          | 65467      | <i>Clytra</i> | <i>quadripunctata</i> | s.str.                |
| 65467         | 65467      | <i>Clytra</i> | <i>quadripunctata</i> | <i>quadripunctata</i> |
| 65468         | 65468      | <i>Clytra</i> | <i>quadripunctata</i> | <i>appendicina</i>    |
| 4794          | 65468      | <i>Clytra</i> | <i>appendicina</i>    |                       |

**Kategorie 8****Automatisierung nicht sinnvoll**

Schließlich bleiben noch diejenigen Fälle, deren korrekte Binärnamen nur von eingearbeiteten Spezialisten zugeordnet werden können.

**Fallbeispiele für Kategorie 8**

*Oreina speciosa* (L. 1767) entspricht *Chrysochloa gloriosa* auct. und *C. vittigera* SUFFRIAN

*Oreina gloriosa* (F. 1781) entspricht *Chrysochloa vittigera* auct. nec SUFFRIAN

Von diesen Arten existieren noch viel mehr Synonyme in zum Teil sehr verwirrenden Kombinationen, vor allem in der englischen Literatur. In der Datenbank ZOBODAT befinden sich zahlreiche Meldungen von Arten mit den Binärnamen *Chrysochloa gloriosa* und *Chrysochloa vittigera*, eingespeichert von E. Geiser vor 1994. Nach der Revision von Kippenberg in FHL 14 hat E. Geiser zahlreiche weitere Exemplare mit dem neuen Schlüssel determiniert. Um deren Daten speichern zu können, muss man für jeden alten und neuen Binärnamen eine andere Nummer vergeben.

Bei der Datenspeicherung ist in diesen Fällen die Supervision bzw. genaueste vorherige Aufbereitung durch Spezialisten notwendig und ebenso bei der Auswertung. Denn ein wesentliches Problem bei der Einspeicherung ist die Tatsache, dass auf den Determinationsetiketten und in der (älteren) Literatur nur z. B. "SUFFRIAN" und eben nicht "auct. nec SUFFRIAN" angegeben ist. Manchmal ist der Autor überhaupt nicht angegeben.

Was mit der Angabe *Oreina gloriosa* tatsächlich gemeint ist, kann man aus folgenden Zusatzinformationen mit mehr oder weniger hoher Wahrscheinlichkeit erschließen: Funddatum (wenn vorhanden), Publikationsjahr, biographischer Daten über die Personen, die das Exemplar gesammelt oder determiniert haben, Kenntnisse über die Sammlung, in der sich das Exemplar befindet, und vor allem genaue Kenntnisse der für bestimmte Zeitabschnitte und Regionen üblichen Bestimmungsliteratur und Checklisten. Z. B. wurden Chrysomeliden in den 50er Jahren im deutschen Sprachraum üblicherweise mit REITTER (1912) bestimmt, Alticinae allerdings auch mit den zahlreichen, verstreut publizierten Tabellen von Heikertinger, die Sammlungsauflistung erfolgte meist nach dem Verzeichnis von HORION (1951).

Auch im folgenden Fall kann nur der – äußert unbefriedigende – status quo festgehalten und mit jeweils unterschiedlichen Nummern versehen werden, ohne dass im Programm ein sinnvoller Zusammenhang hergestellt werden könnte:

*Galerucella sagittariae* GYLLENHAL

*Galerucella nymphaeae* forma *sagittariae* GYLLENHAL nec WEISE

*Galerucella grisescens* (JOANNIS 1866)

= *Galerucella sagittariae* WEISE nec GYLLENHAL

Nach FHL14 sind derzeit folgende Namen gültig, aber keineswegs unumstritten:

*Galerucella nymphaeae* (LINNÉ)

*Galerucella grisescens* (JOANNIS)

*Galerucella sagittariae* (GYLLENHAL)

*Galerucella aquatica* (FOURCROY)

*Galerucella kersteni* (LOHSE)

Die Namensgleichheit verschleiert hier, dass sich die gleichen Namen in beiden Listen auf durchaus verschiedenen Klasseneinteilungen<sup>2</sup> der konkreten Exemplare beziehen können. Eine Zuordnung ist nur mit einer Neudetermination möglich. Dieser Artkomplex ist außerdem bisher erst zum Teil und nur an norddeutschem Material revidiert worden.

## 8. Häufigkeitsanalyse der verschiedenen Fälle

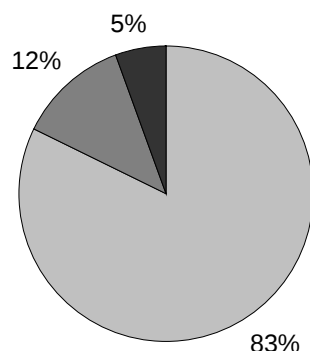
In der folgenden Tabelle sind die Häufigkeiten der vorhin beschriebenen Fälle bezogen auf die Chrysomeliden Mitteleuropas zusammengefasst. Die Zahlen stammen aus der Chrysomeliden-Datei Geiser bzw. aus der Datenbank ZOBODAT. Es sind dabei keineswegs alle Synonyme der Chrysomeliden Mitteleuropas berücksichtigt. Die angegebenen Zahlen sind daher Untergrenzen!

<sup>2</sup> im Sinne der Relationenalgebra, nicht im Sinne der biologischen Systematik oder der Schulverwaltung oder der politischen Ideologie

**Tab. 1:** Häufigkeitsverteilung der Synonymie-Kategorien 1 bis 8 bezogen auf die Anzahl der Arten und der Binärnamen bei mitteleuropäischen Chrysomeliden aus der Datei Geiser. Die Anzahl der Arten, deren Namen unverändert geblieben ist, ist sicher zu hoch. Diese Zahl bezieht sich nur auf Namensänderungen seit 1987.

|                                                              | Anzahl der Arten | Anzahl der Binärnamen |                                                                                  |
|--------------------------------------------------------------|------------------|-----------------------|----------------------------------------------------------------------------------|
| Name unverändert                                             | 366              | 366                   |                                                                                  |
| Kategorie 1<br>umkehrbar eindeutige, einfache Namensänderung | 154              | 330                   | vollständig algorithmierbare Synonymie-Probleme bei 180 Arten mit 376 Binärnamen |
| Kategorie 2<br>verschiedene Schreibweisen                    | 16               | 32                    |                                                                                  |
| Kategorie 3<br>aus Unterarten werden Arten                   | 3 => 6           | 6                     |                                                                                  |
| Kategorie 4<br>Abspaltung einer neuen Unterart               | 1                | 2                     |                                                                                  |
| Kategorie 5<br>Arten werden vereinigt                        | 6 => 3           | 6                     |                                                                                  |
| Kategorie 6<br>Arten werden abgespalten                      | 24 => 44         | 72                    | weitgehend algorithmierbar bei 27 Arten mit 81 Binärnamen                        |
| Kategorie 7<br>aus Arten werden Unterarten                   | 3                | 9                     |                                                                                  |
| Kategorie 8<br>nicht algorithmierbar                         | 12               | > 12                  |                                                                                  |
| Gesamt                                                       | 585              | 835                   |                                                                                  |

Von den Synonymieproblemen sind insgesamt 219 Arten und 469 Binärnamen betroffen. Davon gehören 180 Arten mit 376 Binärnamen zur Kategorie 1 bis 5 und sind somit vollständig algorithmierbar. Bei 27 Arten bzw. 81 Binärnamen kann man die Synonymieprobleme durch die Aggregat-Nummern teilweise algorithmieren. Nicht algorithmierbare Synonymieprobleme treten demnach nur bei (derzeit) mindestens 12 mitteleuropäischen Chrysomelidenarten auf.



**Abb. 1:** Verteilung der Synonymieprobleme auf die verschiedenen Kategorien. Hellgrau: Vollständig algorithmierbar. Dunkelgrau: teilweise algorithmierbar. Schwarz: nicht algorithmierbar.

Die Zahlen beziehen sich auf die Anzahl der Chrysomelidenarten in Mitteleuropa (219), die seit 1987 mindestens ein Synonym erhalten haben.

## 9. Schlussfolgerungen

Wie die vorigen Ausführungen inklusive der Häufigkeitsanalyse zeigen, können die allermeisten Synonymieprobleme mit einem einfachen und praktikablen Algorithmus in einer Biodiversitätsdatenbank gehandhabt werden. Mit den beschriebenen Verfahren ist es weitgehend möglich, die Datenspeicherung und Auswertung Nichtspezialisten zu überlassen. Das Verfahren liefert auch genaue Kriterien, in welchen Fällen bei der Speicherung und Auswertung Spezialisten hinzugezogen werden müssen. Unbedingt notwendig ist die Mitarbeit von Spezialisten bei der Zuordnung konkreter Synonyme zu einer der Kategorien und bei Vergabe der Synonymnummern!

Die Handhabung der Synonymieproblematik ist in den meisten europäischen Datenbanken bisher nur unbefriedigend gelöst. Es ist gewiss kein Zufall, das Finnland die derzeit umfangreichste Evertebraten-Datenbank besitzt. Die Artenzahl und damit auch die Synonymzahl ist dort wesentlich geringer als in Mitteleuropa. ZOBODAT ist immerhin die zweitumfangreichste Evertebraten-Datenbank und hat wesentlich mehr Binärnamen in ihren Grunddatentabellen als die drittstärkste Datenbank in Großbritannien (Institute of Terrestrial Ecology, Monkswood, UK). Auf den britischen Inseln beträgt die Artenzahl der terrestrischen Evertebraten etwa 30 % der entsprechenden Artenzahlen in Mitteleuropa. In ZOBODAT sind 3,1 Millionen artbezogene Datensätze von mehr als 30 000 Arten gespeichert, mit deutlich über 30 000 Binärnamen! Eine praktische Lösung der Synonymieproblematik ist bereits für wesentlich weniger umfangreiche Datenbanken unverzichtbar.

Auch wenn sich mit den Synonymnummern die Systematik in der Datenbank sehr einfach aktualisieren lässt (Änderungen nur an einer einzigen Stelle, einfacher geht es nicht!), muss dringend davor gewarnt werden, die Grunddatentabellen ständig auf dem neuesten systematischen Stand halten zu wollen.

Welcher Name gültig ist und welcher synonym, darüber sind sich oft auch die Spezialisten nicht einig! Aktualisierungen sollten sich an weiterverbreiteten Standardwerken orientieren und nur fallweise von Spezialisten im Zuge einer größeren Revision durchgeführt werden.

- Jede Biodiversitätsdatenbank ist personell stark bis sehr stark unterbesetzt. Es gibt sehr viele Tätigkeiten, die beim Betrieb wesentlich höhere Priorität haben als die Aktualisierung der Systematik.
- Je nach Fragestellung und Verwendungszweck werden bei Auswertungen unterschiedliche Namen gewünscht. Bei einem Fachgutachten im Auftrag der Naturschutzbehörde sollen manchmal die Namen der vorhandenen Lokalfaunistik oder -floristik verwendet werden, bei Auswertungen mit EU-Relevanz und internationaler Kooperation sind die Namen eines verbreiteten europäischen Standardwerkes erwünscht. Ein Spezialist möchte z. B. die aktuellen Namen oder ausdrücklich die älteren Namen für eine Revision.
- Es ist nicht Aufgabe einer Biodiversitätsdatenbank, die Daten jeweils nach dem neuen systematischen Stand auszugeben, sondern sie korrekt und verlustlos abzuspeichern, sodass je nach Auswertungswunsch verschiedene Systeme Verwendung finden können und die Synonyme zusammengehörig ausgewertet werden. Das ist mit unserem Verfahren möglich.

Die meisten Programme, die für verschiedenen Zwecke eingesetzt werden, automatisieren 80 % bis 95 % aller auftretenden Fälle. Mit 95 % entsprechen wir mit unserem Verfahren bereits den strengen Anforderungen der Software-Entwicklung. Ein ganz wesentlicher Vorteil des beschriebenen Verfahrens ist die Möglichkeit, zukünftige Erkenntnisse der Systematik in sehr einfacher Weise zu implementieren. Die Synonymieprobleme der nicht algorithmierbaren Fälle müssen von Systematikern und nicht von Informatikern gelöst werden, die Lösung ist dann algorithmierbar.

Die personellen und finanziellen Mittel sind bei allen Biodiversitätsdatenbanken deutlich geringer, als zur Erfüllung ihrer Aufgaben notwendig wären. Daher ist es sehr wichtig, automatisierbare Prozesse zu automatisieren und zwar so, dass dadurch tatsächlich der Zeitaufwand zur Lösung eines Problems reduziert wird. Mit unserem Verfahren kann ein Großteil der zeitaufwendigen Synonymisierung auf sehr einfache Weise an die Datenbank delegiert werden.

## 10. Danksagung

Die Idee mit den Synonymienummern, die so viele Fälle ebenso einfach wie praktisch abdeckt, geht auf Univ.-Prof. Dr. Ernst Reichl zurück. Wir haben ihm auch das geniale Grundkonzept der ZOBODAT zu verdanken. Man lernt zwar aus eigenen Fehlern (hoffentlich!) und auch aus den Fehlern anderer. Dann weiß man, wie man es nicht machen soll. Aber wie man es am besten machen soll, lernt man an vorbildlichen Leistungen anderer!

Vorbildlich, aber unerreichbar für uns sind die Englischkenntnisse eines Native Speakers, die uns Frau Deborah Fölsche-Forrow, inklusive ihrer langjährigen Erfahrung mit naturwissenschaftlichen Fachartikeln, für unser englisches Abstract zur Verfügung stellte.

Weiters sind wir vielen Kollegen zu Dank verpflichtet, mit denen wir die Probleme der Systematik und der Synonyme in Datenbanken seit über 25 Jahren diskutiert haben. Wir können sie gar nicht alle namentlich aufzählen. Namentlich genannt werden muss allerdings Dr. Remigius Geiser, der die Nomenklaturregeln – sowohl in biologischer als auch in sprachwissenschaftlicher Hinsicht – kennt wie kaum jemand anderer. Ihm verdanken wir neben intensiven Diskussionen zum Thema auch eine Reihe von wesentlichen Hinweisen auf missverständliche Formulierungen. Wir können bei ihm sicher sein, dass er keinen kritisierbaren Aspekt im Manuskript übersehen hat. Eine solch fundierte Kritik vor dem Druck kann sich jeder Autor nur wünschen – wenngleich wir ihm nicht in jedem Punkt gefolgt sind!

## 11. Zusammenfassung

Synonyme, also die unterschiedlichen wissenschaftlichen Namen für ein und dieselbe Art, verursachen bei artbezogener Datenspeicherung und Auswertung in Biodiversitätsdatenbanken erhebliche Probleme. Die Autoren haben ein Verfahren ausgearbeitet und in der Datenbank ZOBODAT implementiert, das die Zuordnung von Synonymen zum jeweils gültigen Namen ganz oder teilweise automatisiert und dennoch eine korrekte artbezogene Datenspeicherung und Auswertung mit minimalem Aufwand gewährleistet.

Zunächst wurde in einer Anforderungsanalyse erhoben, welche Kriterien eine Lösung der Synonymieproblematik in Biodiversitätsdatenbanken erfüllen muss. Wesentliche Kriterien sind dabei

- Einspeicherungsmöglichkeit unter dem aktuellen und unter einem früheren Namen.
- Problemlose Erweiterbarkeit bei künftigen Namensänderungen.
- Artbezogene Auswertung durch alle Synonymieketten hindurch, ohne Informationsverlust.

- Minimaler Programmier- und Wartungsaufwand der Datenbank.
- Der Einspeicherungsmodus muss praktikabel sein und von Nichtspezialisten durchgeführt werden können.

Am Beispiel der Chrysomeliden Mitteleuropas wurden die Synonymieprobleme 8 verschiedenen Kategorien zugeordnet. Für die Kategorien 1 bis 7 wird eine Lösung in der Datenbank vorgestellt und die Erfüllung der Anforderungen überprüft. Die Kategorie 8 enthält die Synonymieprobleme, bei denen eine Automatisierung beim derzeitigen Kenntnisstand der Systematik nicht sinnvoll ist und die von Spezialisten der Gruppe eingespeichert bzw. dafür vorbereitet werden müssen.

#### **Lösungsansatz für Kategorie 1 bis 5**

Jeder wissenschaftliche Name enthält eine Nummer, unter der diese Artbezeichnung eindeutig identifizierbar ist. Weiters wird jedem wissenschaftlichen Namen eine Synonymienummer zugeordnet. Bei einem gültigen Namen sind die Nummer der Artbezeichnung und die Synonymienummer identisch. Bei einem Synonym erhält die Synonymienummer den Zahlwert der Identifikationsnummer des gültigen Namens zugeordnet.

#### **Lösungsansatz für Kategorie 6 bis 7**

In diese Kategorien fallen Abspaltungen von neuen Arten, wobei die "Stammart" nomenklatorisch unverändert bleibt. Die neu abgespaltene Art kann einfach durch eine neue Identifikationsnummer in der Datenbank gekennzeichnet werden. Die bereits unter dem Namen der "Stammart" eingespeicherten Datensätze beziehen sich aber nun auf zwei Arten, die noch nicht getrennt wurden. Deren Identifikationsnummer wird in der Artentabelle als "Aggregat" gekennzeichnet. Um auch Datensätze der Art in der neuen Bedeutung, also ohne Exemplare der abgespaltenen Art, speichern zu können, erhält der Binärname der Stammart eine weitere Nummer und wird als *sensu stricto* gekennzeichnet. In einer weiteren Tabelle wird festgehalten, dass diese drei voneinander unabhängigen Nummern zusammengehören.

Wie die Nummernvergabe in der Praxis bei den verschiedenen Synonymiekategorien zu erfolgen hat, wird mit Beispielen mitteleuropäischer Chrysomelidenarten erläutert. Die Häufigkeitsverteilung der Synonymieprobleme bezogen auf die Artenzahl ergibt, dass bei 84 % der Arten mit Synonymen deren Handhabung vollständig algorithmierbar und bei weiteren 10 % immerhin teilweise algorithmierbar ist. Diese Zahlenverhältnisse treffen in der Größenordnung auf die meisten mitteleuropäischen Insektenarten zu, vor allem auf Käfer und Schmetterlinge, die einen überproportional hohen Anteil an Evertebratendaten in Biodiversitätsdatenbanken ausmachen. Das Verfahren ist somit eine sehr brauchbare Methode, die Datenspeicherung und Auswertung mit minimalem Aufwand und ohne Informationsverlust zu handhaben. Ein weiterer Vorteil ist, dass die Datenspeicherung und Auswertung weitgehend von Nichtspezialisten durchgeführt werden kann. Das Verfahren liefert auch klare Kriterien, für welche Schritte und in welchen Fällen man Spezialisten direkt oder für die Supervision braucht.

## **12. Literatur**

### **FHL 1-15 bedeutet:**

- FREUDE H., HARDE K.W. & G.A. LOHSE (1964-1983) (Hrsg.): Die Käfer Mitteleuropas. Band **1-11**. — Krefeld: Goecke & Evers.
- LOHSE G.A. & W. LUCHT (1989 – 1994) (Hrsg.): Die Käfer Mitteleuropas. Bd. **12-14** (1.-3. Supplementband). — Krefeld: Goecke & Evers.
- LUCHT W. & B. KLAUSNITZER (1998) (Hrsg.): Die Käfer Mitteleuropas. Bd. **15** (4. Supplementband). — Goecke & Evers im Gustav Fischer Verlag.

- FUSS G., GEISER E. & R. PATZNER (2005): On the Host Plants of Several Leaf Beetles of Central Europe – the Problem of Fame and Evidence (Coleoptera: Chrysomelidae). — *Koleopterologische Rundschau* 75: 359-371.
- GEISER E. (1996): Probleme der Datenerfassung, Determination und Nomenklatur bei umfangreichen faunistischen Datenbanken. — Verhandlungen des XIV. Internationalen Symposiums über Entomofaunistik in Mitteleuropa (SIEEC), 4.-9. September 1994, München: 344-347.
- GEISER E. (1998): Die tiergeographische Datenbank ZOODAT, eine wertvolle Datenbasis für die Biodiversitätsforschung. — *Stapfia (Linz)* 55: 35-46.
- GEISER E. (2001): Die Käfer des Landes Salzburg. Faunistische Bestandserfassung und tiergeographische Interpretation. — *Monographs on Coleoptera (Wien)*, Vol. 2.
- GEISER E. (2004): Chrysomelidae (Insecta: Coleoptera). — Checklisten der Fauna Österreichs, Nr. 1. Biosystematics and Ecology Series, Nr. 22 (Österreichische Akademie der Wissenschaften, Wien): 1-30.
- GEISER E. (2005): Die entscheidenden Faktoren für den langfristigen Erfolg einer tiergeographischen Datenbank – am Beispiel der ZOBODAT (früher: ZOODAT). — *Linzer biologische Beiträge* 37 (1): 49-69.
- GROSS J., FATOUROS N.E. & M. HILKER (2004): The significance of bottom-up effects for host plant specialization in *Chrysomela* leaf beetles. — *OIKOS* 105: 368-376.
- HORION A. (1951): Verzeichnis der Käfer Mitteleuropas. — Stuttgart: Verlag A. Kernen.
- LUCHT W. (1987): Die Käfer Mitteleuropas. Katalog. — Krefeld: Goecke & Evers.
- MALICKY M. (2001): Die biogeographische Datenbank ZOBODAT (ehemalige ZOODAT) am OÖ. Landesmuseum / Biologiezentrum Linz, Oberösterreich. – eine öffentliche Einrichtung für Wissenschaft und Naturschutz im Umfeld nationaler und internationaler Kooperationsmöglichkeiten. — *Entomologica Austriaca* 4: 8-14.
- REITTER E. (1912): Fauna Germanica, Band IV. — Stuttgart: K. G. Lutz Verlag.
- EIS.net: [www.eis-international.org/index.php](http://www.eis-international.org/index.php)
- GBIF: [www.gbif.org](http://www.gbif.org)
- ZOBODAT: [www.zobodat.at](http://www.zobodat.at)

Anschrift der Verfasser: Dr. Elisabeth GEISER  
St.-Julien-Straße 2 /314  
5020 Salzburg, Austria  
E-Mail: [Elisabeth.Geiser@gmx.at](mailto:Elisabeth.Geiser@gmx.at)

DI Michael MALICKY  
Biologiezentrum der Oberösterreichischen Landesmuseen  
J.-W.-Klein-Straße 73  
4040 Linz, AUSTRIA  
E-Mail: [m.malicky@landesmuseum.at](mailto:m.malicky@landesmuseum.at)

Mag. Dr. Elisabeth GEISER, Jahrgang 1952, Studium der Mathematik, Biologie und Geowissenschaften, Promotion in Zoologie, derzeit von Beruf Software-Engineer in der Entwicklungsabteilung einer innovativen Elektronikfirma. Seit über 25 Jahren Mitarbeit bei ZOBODAT. Im Rahmen dieser Zusammenarbeit verfasste sie – unter anderem – eine umfangreiche Faunistik der Käfer des Landes Salzburg und eine Checkliste der Chrysomeliden Österreichs.

DI Michael MALICKY, Jahrgang 1969, Studium der Informatik, seit 1990 Mitarbeiter und seit 1999 Leiter von ZOBODAT (eine der größten europäischen Datenbanken, insbesondere für Evertrebraten) am Biologiezentrum in Linz. Mitarbeit bei zahlreichen EU- und anderen internationalen Projekten über Biodiversitätsdatenbanken, technischer Koordinator von GBIF-Österreich.

# ZOBODAT - [www.zobodat.at](http://www.zobodat.at)

Zoologisch-Botanische Datenbank/Zoological-Botanical Database

Digitale Literatur/Digital Literature

Zeitschrift/Journal: [Entomologica Austriaca](#)

Jahr/Year: 2007

Band/Volume: [0014](#)

Autor(en)/Author(s): Geiser Elisabeth, Malicky Michael

Artikel/Article: [Zur Algorithmierbarkeit von Synonymieproblemen in Biodiversitätsdatenbanken. Anforderungsanalyse und praktische Lösungsvorschläge am Beispiel der Chrysomeliden \(Insecta: Coleoptera\) 17-42](#)