

## Probleme der Datenerfassung, Determination und Nomenklatur bei umfangreichen faunistischen Datenbanken

Elisabeth GEISER

Faunistische Datenbanken schießen heute wie Schwammerl aus dem Boden. Neben den großen nationalen (z.B.: Tiergeographische Datenbank Österreichs ZOODAT oder das Schweizerische Zentrum für die kartographische Erfassung der Fauna SZKF) gibt es heute gerade bei Insekten eine Fülle von regionalen und privaten faunistischen Datenbanken.

Wer sich mit Fragen der Tiergeographie oder des Artenschutzes befaßt, ist auf faunistische Daten von größeren Flächen angewiesen. Die Fundmeldungen in den verschiedenen regionalen Datenbanken sollten daher "kompatibel" sein, d.h. sich automatisch-elektronisch zu einer größeren Datenbank zusammenfassen lassen, von der einem dann – im Idealfall – der Computer eine schöne lange Liste mit allen eingespeicherten Fundmeldungen und einer schönen Karte ausdrückt.

Informationstheoretisch ist das möglich. Warum es in der Praxis fast nie klappt, möchte ich Ihnen hier erläutern. Meine folgenden Ansichten zu diesem Problemkreis sind das Ergebnis meiner 14jährigen Mitarbeit bei der Tiergeographischen Datenbank Österreichs ZOODAT.

Bei tiergeographischen Datenbanken können folgende Extremfälle auftreten:

**Fall 1:** Ein Spezialist, der seit Jahren eine schwierige Käfergattung bearbeitet, hat eine private Datenbank von allen Käfern seiner Sammlung angelegt, und allen von ihm jemals untersuchten Exemplaren. Diese Käfer sind daher äußerst zuverlässig bestimmt, niemand kann es besser als er, denn es gibt derzeit gar keinen zweiten Spezialisten dieser Gattung. Der Zuverlässigkeitsgrad der Determinationen seiner Datenbank ist daher > 95 %.

Derselbe Spezialist verwendet in seiner Datenbank die neueste Nomenklatur, die zum größten Teil auf seinen eigenen Publikationen beruht.

Weil aber

- a) über diese neueste Nomenklatur kein anderer genauso gut Bescheid weiß und
- b) sich kein anderer die (vergebliche!) Mühe macht, laufend so viele Namen in seiner Datenbank zu ändern,

geht die Kompatibilität dieser so zuverlässigen Datenbank gegen 0, d.h. man kann sie mit keiner anderen Datenbank automatisch zusammenkombinieren, ohne schlimme Fehlergebnisse zu produzieren.

**Fall 2:** Das andere Extrem ist frei erfunden, Ähnlichkeiten mit real existierenden Datenbanken sind rein zufällig.

Eine überregionale Umweltschutzbehörde, (z.B. von einem österreichischen Bundes-

land) möchte eine faunistische Datenbank für Naturschutzzwecke anlegen. Um in kurzer Zeit möglichst viele Daten zu bekommen (denn nur so kann man das Vorhaben vor den Geldgebern rechtfertigen), rafft man alles an Daten zusammen, was man kriegen kann.

1. Man schickt eine Hilfskraft in das örtliche Landesmuseum, die an der dortigen Insektenammlung alle Etiketten abschreibt. Diese Hilfskraft hat weder eine besondere Qualifikation dafür noch wurde sie von jemandem angeleitet, der was davon versteht, da man annimmt: abschreiben kann schließlich jeder, und was in der Landessammlung steckt, muß ja stimmen!
2. Man beschäftigt mittels Werkvertrag einen Zoologiestudenten, der innerhalb weniger Monate die gesamte(!) faunistische Literatur dieses Landes zusammenstellen soll. Diese Daten werden dann wieder von einer Hilfskraft eingegeben.
3. Man speichert alle Funddaten ein, die von diversen ökologischen Gutachten, aus Exkursionsprotokollen und aus den Diplomarbeiten der örtlichen Universität stammen.

Die örtlich ansässigen Spezialisten werden nicht eingebunden. Zum Teil sind sie den Beamten in der Naturschutzbehörde gar nicht bekannt, zum Teil sind die Spezialisten auf die Naturschutzbehörde sauer, weil diese einerseits den Spezialisten durch Sammelbeschränkungen das Leben schwermacht, andererseits aber umsonst vollständige Listen ihrer Funddaten haben will (ohne finanzielle Abgeltung natürlich), die sie in jahrzehntelanger Mühe zusammengetragen haben.

Da sehr bald klar wird, daß die Insekten in der Museumssammlung, in der Literatur und in den Diplomarbeiten unterschiedliche Namen haben, einigt man sich, als Namen in der Datenbank den des jeweils neuesten umfassenden Standardwerkes einer Tiergruppe zu verwenden, also bei Käfern den Lucht-Katalog von 1987 nach Freude-Harde-Lohse, Band 2 – 11, und bei den Schmetterlingen Forster-Wohlfahrt. Mit der Synonymiesuche wird wiederum unser Werkstudent beauftragt. Er ist Spezialist für Libellen, also versteht er was von Insekten und Nomenklatur (jedenfalls mehr als irgend jemand in der Behörde!) und wird sicher bei allen Gruppen das richtige Synonym herausfinden!

Die Kompatibilität dieser Datenbank ist durch die Verwendung der Standardwerke recht hoch, sagen wir > 90 %.

Die Zuverlässigkeit der so erfaßten Daten geht gegen 0!

In diesem Spannungsfeld zwischen diesen beiden Extremen bewegen sich tiergeographische Datenbanken!

Daraus sieht man folgendes:

Es ist keineswegs sinnvoll

1. alle Daten aufzunehmen, die man kriegen kann,
2. anzustreben, immer die neueste Nomenklatur zu verwenden, und
3. mit ungelerten Hilfskräften statt mit Spezialisten zusammenzuarbeiten.

Welche Daten soll man nun aufnehmen?

Dazu brauchen Sie erstens Leute, die die Zuverlässigkeit von Sammlungen und Grauliteraturangaben beurteilen können. Das kann man bei Sammlungen durch stichprobenartige Überprüfung der Determination bei den Tieren der eigenen Spezialgruppe überprüfen, bei Literatur aller Graustufen mittels Durchsicht auf Plausibilität. (Enthalten die Angaben auffällige Häufungen von seltenen Insekten? Werden Insekten genannt, die in diesem Gebiet höchstwahrscheinlich gar nicht vorkommen? Sind die häufigen, sicher dort vorkommenden Arten auch vertreten? Sind etwa die Käfer ausschließlich nach dem Kosmos-Buch von

Bechyne: "Welcher Käfer ist das?" bestimmt worden usw.) Ganz wesentlich ist auch, wer die Tiere bestimmt hat. Ich selbst kann Staphyliniden-Determinationen nicht direkt auf ihre Zuverlässigkeit nachprüfen. Aber wenn beim Exkursionprotokoll der Universität Gießen in die Hohen Tauern drinnensteht: Staphylinidae det. PUTHZ, dann weiß ich, daß ich eine ganz wertvolle Datenquelle vor mir habe.

Nach dieser Vorauswahl haben Sie auch noch reichlich Daten zur Verfügung! Ein paar fehlerhafte Daten kommen trotzdem mit hinein, das kann man nie 100 %ig ausschließen, auch nicht bei der Spezialistensammlung. Diese wenigen Fehler sind aber nicht so tragisch, denn:

- a) entweder liegt die fehlerhafte Fundmeldung innerhalb des normalen Verbreitungsgebietes, dann stört sie nicht bei der Auswertung.
- b) sie liegt außerhalb des Verbreitungsgebietes. Dann fällt sie sofort auf und man kann der Sache nachgehen. Stellt sich dann heraus, daß es sich wirklich um eine Fehleingabe oder -angabe handelt, kann man diese Fundmeldung aus der Datenbank eliminieren. Manchmal stellt sich auch heraus, daß der abweichende Fundpunkt stimmt, dann hat man neue Erkenntnisse über das Verbreitungsgebiet gewonnen.

Die Datenzuverlässigkeit ist ein relativ leicht zu lösendes Problem, wenn man mit Leuten zusammenarbeitet, die etwas davon verstehen.

Leider gilt das nicht für das zweite große Problem, die Nomenklatur:

Sie alle wissen aus leidvoller Erfahrung, daß sich die "richtigen" Namen von Tieren und Pflanzen ständig ändern, und zwar besonders bei den mitteleuropäischen Insekten.

Allein bei den Käfern Mitteleuropas haben sich seit dem Erscheinen des Standardwerkes von Freude-Harde-Lohse, dessen letzter Käferband 1987 herauskam, in den Nachtragsbänden (1989 bis 1994) mehr als ein Drittel aller Käfernamen geändert, bei manchen Familien sogar bis 80 %! Diese Nachträge sind schon wieder veraltet und der Nachtrag der Nachträge ist bereits in Arbeit. Nur ein Teil dieser Namensänderungen ist auf neue systematische Erkenntnisse zurückzuführen, in vielen Fällen handelt es sich nur um rein archäoliterarische Änderungen.

Diese Namensinstabilität verursacht zahlreiche Probleme, besonders zu leiden haben darunter die tiergeographischen Datenbanken, die oft mehrere 10.000 Artnamen enthalten. Hier nomenklatorisch immer auf dem letzten Stand sein zu wollen, ist nicht durchführbar und auch gar nicht sinnvoll. Sie können zwar die rein formalen Namensänderungen in EDV vollziehen, aber da Sie damit überhaupt keine neuen Erkenntnisse gewinnen, ist das eine völlig nutzlose Mühe! Bei den Namensänderungen, die auf neuen systematischen Erkenntnissen beruhen, sind häufig Arten neu aufgespalten oder zusammengefaßt worden. Hier können Sie überhaupt nicht feststellen, ob die ältere Angabe zu der oder zu jener Art gehört, Sie müssen jedes einzelne Exemplar nachbestimmen. [Wem nicht sofort ein Beispiel bei seiner eigenen Spezialgruppe einfällt: *Chrysochloa vittigera* (SUFFR.) nach MOHR 1966 in FHL), jetzt nach Kippenberg 1994: *Oreina gloriosa* (*vittigera* auct., auch MOHR 1966 in Band 9, nec SUFFR.) und dafür heißt die ehemalige *Chrysochloa gloriosa* (F.) jetzt *Oreina speciosa* (*gloriosa* auct. auch MOHR 1966 in Band 9, *vittigera* (SUFFR.)), wobei die Bearbeitung von Kippenberg eine wesentliche Verbesserung der reichlich verworrenen Verhältnisse bei den *Chrysochloen* darstellt!).

Es ist aber gar nicht die Aufgabe einer tiergeographischen Datenbank, praktisch abrufbereit neueste systematische Kenntnisse zu speichern! Die tiergeographische Datenbank sammelt Daten, damit sie anschließend ein Spezialist bearbeiten kann und nimmt dem Spezialisten dabei viele Vorarbeiten ab. Jetzt weiß er, wo die fraglichen, noch zu revidie-

renden Exemplare sind und kann sie, so er es für notwendig hält, aufsuchen und nachbestimmen.

Für tiergeographische Datenbanken ist es daher am besten, sich in ihrer Nomenklatur nach dem letzten großen Standardwerk der jeweiligen Gruppe zu richten. Das ist bei den Käfern Freude-Harde-Lohse Band 2 bis 11, wobei diejenigen Käfer, die inzwischen echt neu dazugekommen sind, sei es durch Neufunde für Mitteleuropa oder Neubeschreibungen, zusätzlich aufgenommen werden können. Bei neuen Artaufteilungen oder bei Typenrevisionen müssen bei Neubestimmungen Autor und Nomenklatur (z.B. nach KIPPENBERG 1994 in FHL, Band 14) bei den Datenquellen angegeben werden.

Rein formale Namensänderungen, über deren Gültigkeit sich ja oft nicht einmal die jeweiligen Spezialisten einig sind, sind bei tiergeographischen Datenbanken kontraproduktiv. Der Computer soll uns ja die Arbeit erleichtern und nicht die Probleme, die wir in der Systematik und Faunistik ohnehin schon mit der Nomenklatur haben, noch potenzieren.

Mit dieser pragmatischen Vorgangsweise sind zwar nicht alle Probleme gelöst, aber damit bleibt die wertvolle Information, die die großen tiergeographischen Datenbanken enthalten, weiterhin benutzbar. Es würde mich nun sehr interessieren, wie vor allem die nomenklatorischen Probleme von den anderen, umfangreichen Datenbanken gelöst werden und stelle meinen Vortrag hiermit zur Diskussion.

ANSCHRIFT DER VERFASSERIN:

Elisabeth Geiser  
Saint-Julien-Straße 2/314  
A-5020 Salzburg (Österreich)