

Inventare der Artenvielfalt und Urheberrecht *Probleme bei der Erstellung elektronischer Datenbanken*

Von Donat Agosti*

Nach wie vor ist nicht klar, wie viele Arten es überhaupt gibt. Hauptverantwortlich dafür ist die Schwierigkeit, die Vielfalt der Lebewesen in der Welt zu erfassen. Eine neue Barriere für den Aufbau moderner Datenbanken stellt auch das Urheberrecht dar.

Seit Ende der achtziger Jahre spricht man von einer Biodiversitätskrise, und spätestens seit dem Erdgipfel in Rio de Janeiro von 1992 besteht ein Konsens darüber, dass Veränderungen und Eingriffe in die Natur im Rahmen einer sogenannten nachhaltigen Nutzung erfolgen sollten. Doch wie von Online-Veröffentlichungen öffnet sich ein neuer Informationsgraben zwischen reichen und ärmeren Institutionen, speziell solchen der Dritten Welt, welche die exorbitanten Kosten der notwendigen, oft nur als Multipack verkauften Zugriffsgenehmigungen für wissenschaftliche Zeitschriften nicht bezahlen können.

Erstmals spürbar sind die Probleme vor allem bei den Ameisen geworden. Bei ihnen und den Fischen ist der digitale Zugang zur Artenvielfalt und zur dazugehörenden Literatur am weitesten fortgeschritten. Bei den Ameisen liegt mittlerweile ein Katalog vor, der Ende 2001 bereits über 11 000 bekannte Arten und entsprechende wissenschaftliche Arbeiten umfasste.

Man geht davon aus, dass die Beschreibungen von allen Ameisenarten digital erfasst sind, ebenso auch mindestens 90 Prozent der weiteren Literatur. Insgesamt handelt es sich um 3500 Arbeiten, die Beschreibungen von Arten enthalten, zusätzlich 5400 Arbeiten über deren Verbreitung, Morphologie oder Genetik. Etwa 24 000 weitere Publikationen befassen sich mit Verhalten und angewandten Gebieten sowie der Bekämpfung von Ameisen. Bis heute sind davon 1000 Arbeiten als Bilddokumente (in PDF-Format) gespeichert und mit den entsprechenden Angaben in der Online-Datenbank verknüpft. Ein laufendes, durch die Smithsonian Institution unterstütztes Projekt dürfte bis Ende 2003 alle 3500 Arbeiten mit Artenbeschreibungen online zugänglich machen. Man geht im Durchschnitt von 23 Seiten pro Arbeit aus. Das entspricht insgesamt 80 500 Seiten. Für jede Seite wird mit Kosten von etwa einem Dollar und durchschnittlich 80 Kilobyte gerechnet, total also mit 6,4 Gigabyte Speicherkapazität und Aufwendungen von 80 500 Dollar.

Technisch stellt dies alles kein Problem dar. Schwierigkeiten gibt es aber mit dem Urheberrecht und beim Auffinden der Arbeiten. Zurzeit gib es weltweit höchstens vier Bibliotheken, in denen alle 2600 Arbeiten, in denen erstmals eine Art beschrieben wurde, vorliegen. Diese sind von 446 verschiedenen Autoren verfasst und in 511 Zeitschriften und 103 Büchern über alle Kontinente hinweg veröffentlicht worden. Das älteste Dokument stammt von Carolus Linnaeus aus dem Jahr 1758. Um die Arbeiten auf dem Internet zugänglich zu machen, muss nun aber eine Bewilligung von den Autoren und den Verlagen eingeholt werden. Dies wird für alle neueren Arbeiten gemacht, besonders für solche, die weniger als 75 Jahre alt sind. Bei älteren Werken, so eine Faustregel, kann man sich den Aufwand sparen, für jedes Ursprungsland das nationale Urheberrecht zu studieren.

Kommerzielle Interessen der Verlage

Zwar sind nur ganz wenige Autoren dagegen, ihre Arbeiten online zugänglich zu machen. Und bei diesen handelt es sich ausnahmslos um Wissenschaftler der entwickelten Länder. Auch alle

befragten wissenschaftlichen Vereine, welche Zeitschriften verlegen, äusserten sich zustimmend. Im Gegensatz dazu verweigern aber die meisten kommerziellen Verleger die Freigabe, und zwar unabhängig davon, ob sie ihre Arbeiten online anbieten oder nicht, und selbst in den Fällen, in

Erschwerung der Forschung

Da es sich zum grossen Teil um Arten aus den Tropen handelt, verstösst diese Praxis auch gegen den Artikel 17 der Uno-Konvention zur Artenvielfalt. Dieser schreibt vor, dass Daten, die aus der Erforschung der Artenvielfalt gewonnen werden, wieder ins Ursprungsland zurückfliessen müssen. Zudem behindert die Blockierung des freien Zugangs zu allen Publikationen das Entstehen neuer Forschungsergebnisse und -gebiete. Zurzeit werden nämlich – ähnlich wie in den Genomics – Methoden entwickelt, um nach Zusammenhängen innerhalb der enormen Menge von Daten zu suchen, die in den Beschreibungen von Morphologie und Verhalten oder auch der Verbreitung enthalten sind. Auch möchte man Daten über die Verbreitung extrahieren, um für vorgegebene Orte die Artenvielfalt schneller berechnen zu können. Dafür werden Prototypen-Programme entwickelt, die auf Texterkennung und XML-Technologie zur logischen Markierung des Textes basieren. Sie bilden eine der Grundlagen für die kürzlich gegründete «Global Biodiversity Information Facility». Deren Zukunft ist aber nun durch eine zunehmend restriktivere Politik im Bereich des Urheberrechts in Frage gestellt.

In der allgemeinen Biodiversitätsforschung ist man sich des Problems des Datenzugangs allerdings noch kaum bewusst, da überhaupt erst seit kurzem zaghafte Schritte ins digitale Zeitalter unternommen werden. Dass das Verhalten der Verlage tatsächlich ein Problem sein kann, zeigt sich aber nicht nur bei den Arbeiten über Ameisen, sondern auch beispielsweise bei PubMed Central, einem amerikanischen Konsortium, dessen Online-Archiv Zugang zu medizinischen wissenschaftlichen Arbeiten bietet, die älter als ein Jahr sind. Trotz dem grossen Interesse der Wissenschaftler beteiligen sich die Verleger nur sehr zurückhaltend.

Alternative Modelle

Der Amerikaner Tom Moritz hat nun unter dem Namen «Biodiversity Commons» kürzlich ein neues Modell für die möglichst effiziente Nutzung riesiger Datenmengen in der Forschung vorgeschlagen. Die Idee ist, dass nur Institutionen an solchen Projekten teilnehmen können, die sich dazu verpflichten, ihre Daten für Forschungszwecke frei zur Verfügung zu stellen. Ein anderes Modell stammt von der wissenschaftlichen Zeitschrift «Nature», die alle Beschreibungen von Arten innerhalb ihrer Seiten via einen offiziellen Depositar frei zugänglich macht, dabei geben auch die Autoren nach der Veröffentlichung ihre Texte frei. Eine solche Lösung schliesst eine parallele kommerzielle Nutzung nicht aus. Als Alternative könnte auch in Betracht gezogen werden, dass Beschreibungen nur noch in Zeitschriften publiziert werden, die kein Urheberrecht kennen – eine Lösung, die längerfristig durch die sukzessive ins Rollen kommende Diskussion über Online-Beschreibungen ohnehin Realität werden könnte.

* Der Autor ist Biologe und massgeblich am Aufbau der Datenbank über Ameisen beteiligt (<http://antbase.org>).