

Das Projekt MoveVRE - Movebank Virtual Research Environment (Virtuelle Forschungsumgebung)

Petra Hätscher

22

Welche Routen fliegen Störche nach Süden? Wo pausieren sie? Welchen Einfluss haben Wind und Wetter auf die Flugrouten? Diese und viele andere Fragen rund um Tierbewegungen beschäftigen die Ornithologen und andere mit Tierbewegungen befassten Biologen seit Beginn der Forschung in diesen Themenfeldern. Doch noch nie war es so einfach wie heute, große Mengen an Daten zu erheben und zu verarbeiten. Für das Thema Tierbewegungen existiert die Datenbank Movebank (movebank.org), in der von Wissenschaftlern und Wissenschaftlerinnen rund um die Welt Daten ihrer Forschung eingespielt und anderen zur Verfügung gestellt werden. Dadurch entsteht ein sehr großer Datenpool, der bei weiteren Forschungen zu den entsprechenden Spezies als Grundlage genutzt werden könnte.

Movebank wurde von Prof. Dr. Martin Wikelski, Direktor des Max Planck Institutes für Ornithologie (Standort Radolfzell) und Lehrstuhlinhaber am Fachbereich Biologie der Universität Konstanz, entwickelt und betrieben. Im Jahr 2009 bildete sich an der Universität Konstanz ein Konsortium, um sich auf eine Ausschreibung der Deutschen Forschungsgemeinschaft (DFG) zum Thema Virtuelle Forschungsumgebungen zu bewerben. Dem Konsortium gehörten neben der AG Wikelski die Bibliothek, das Rechenzentrum und der Lehrstuhl von Prof. Dr. Daniel Keim am Fachbereich Informatik (Schwerpunkt Datenanalyse und Visualisierung) an.

Was ist eine Virtuelle Forschungsumgebung? Die entsprechende AG der Schwerpunktinitiative „Digitale Information“ der Allianz der Wissenschaftsorganisationen (Allianzinitiative) definiert den Begriff folgendermaßen: „Eine Virtuelle Forschungsumgebung (Virtual Research Environment - VRE) ist eine Arbeitsplattform, die eine kooperative Forschungstätigkeit durch mehrere Wissenschaftlerinnen und Wissenschaftler an unterschiedlichen Orten zu gleicher Zeit ohne Einschränkungen ermöglicht. Inhaltlich unterstützt sie potentiell den gesamten Forschungsprozess – von der Erhebung, der Diskussion und weiteren Bearbeitung der Daten bis zur Publikation der Ergebnisse - während sie technologisch vor allem auf Softwarediensten und Kommunikationsnetzwer-

ken basiert. Virtuelle Forschungsumgebungen sind wesentliche Komponenten moderner Forschungsinfrastrukturen und spielen eine entscheidende Rolle für die Produktivität und Wettbewerbsfähigkeit der Forschung.“¹ Die DFG-Ausschreibung enthielt als eine unabdingbare Voraussetzung für eine erfolgreiche Antragstellung die Zusammenarbeit von wissenschaftlichen Arbeitsgruppen mit Infrastrukturanrichtungen wie Bibliotheken und Rechenzentren. Damit sollte der nachhaltige Betrieb sowie die Berücksichtigung von Erfahrungen aus jeweils anderen Wissenschaftsdisziplinen gewährleistet werden. Das Projekt hatte eine Laufzeit von April 2010 bis Dezember 2012. Drei Arbeitspakete wurden definiert:

- Aufbau eines Datenrepositoriums für die dauerhafte, unveränderbare Speicherung von Datensätzen zu Tierbewegungen, versehen mit jeweils einem Persistent Identifier. (Hauptverantwortlich: Bibliothek)
- Einbindung von Werkzeugen zur Zusammenarbeit für Wissenschaftlerinnen und Wissenschaftler in die Datenbank Movebank. (Hauptverantwortlich: Rechenzentrum)
- Entwicklung verbesserter Visualisierungsmöglichkeiten für die Daten. (Hauptverantwortlich: AG Keim)

Da die Bibliothek neben der Projektleitung und -koordination vor allem für das Themenfeld 1 - Aufbau eines Datenrepositorium - verantwortlich zeichnete und das ein Thema ist, das Bibliotheken auch in Zukunft beschäftigen wird, werden diese Ergebnisse näher beschrieben. Themenfeld 2 (Werkzeuge zur Zusammenarbeit - collaboration tools) wird zu einem späteren Zeitpunkt nochmal aufgegriffen, da es in den kommenden Jahren weiter Thema bleiben wird. Themenfeld 3 (Visualisierung großer Datenmengen) wird hier durch ein Bild verdeutlicht, bei dem die Flugrouten in Abhängigkeit zur Windgeschwindigkeit für Störche gezeigt werden. Die Abhängigkeiten sind schneller erkennbar als bei einer nicht visualisierten Darstellung, z. B. bei einer Auflistung der Daten in Tabellenform.

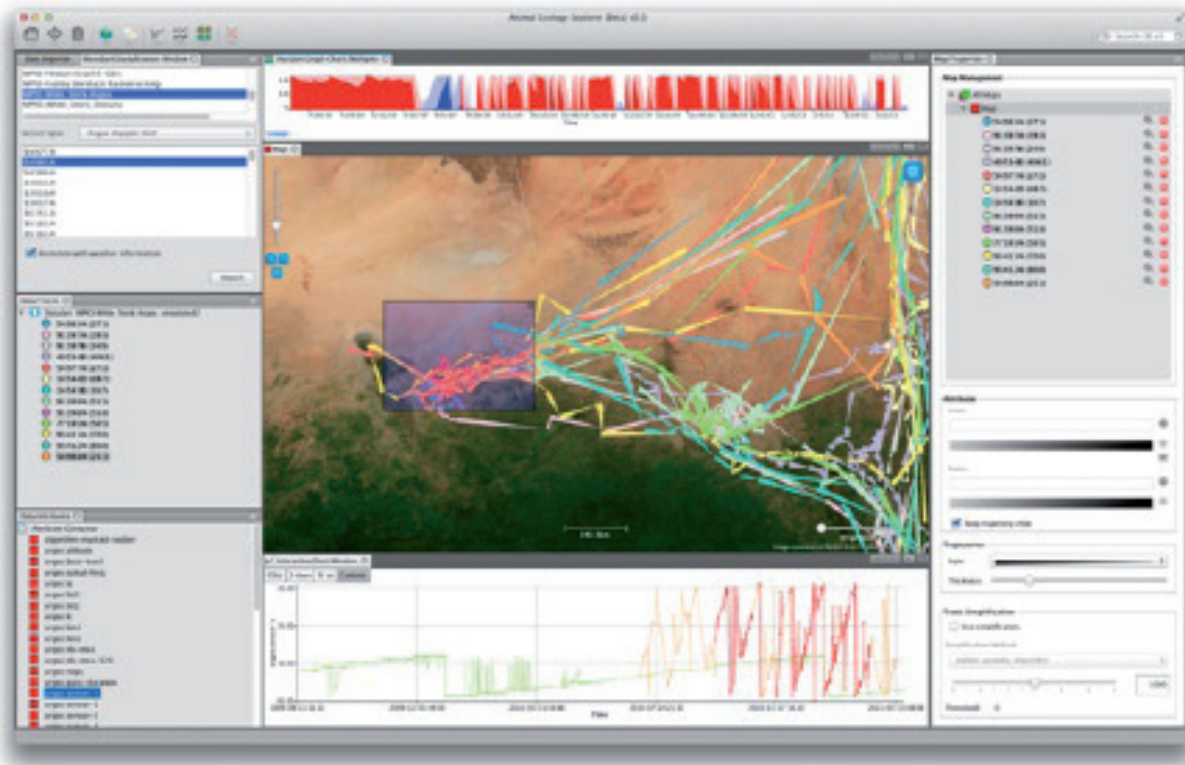


Abb. 1: Visualisierung von Vogelfluglinien in Abhängigkeit von weiteren Kriterien (z.B. Windgeschwindigkeit)

In der Datenbank Movebank werden die von den Wissenschaftlerinnen und Wissenschaftlern erhobenen Forschungsprimärdaten direkt eingegeben. Sie können in der Datenbank durch Berechtigte, in der Regel die Forschenden selbst, verändert und gelöscht werden. Es ist eine „lebendige“ Datenbank in dem Sinne, dass sie vor allem als Arbeitsinstrument dient. Langzeitsicherung, dauerhafte Zitierfähigkeit und dauerhafte Präsentation spielen eine untergeordnete Rolle.

Diese Lücke schließt das die Datenbank Movebank ergänzende Datenrepositorium (Movebank Data Repository)². Im Datenrepositorium werden Forschungsdatensätze dauerhaft gespeichert. Sie dürfen nicht mehr verändert werden und erhalten für die Zitierbarkeit einen Persistent Identifier, eine DOI (Digital Object Identifier), wie sie auch aus der der Publikationswelt bekannt ist. Jeder gespeicherte Datensatz wird mit einer DOI ausgezeichnet und wird dadurch zitierbar, kann einfach mit der entsprechenden Publikation verknüpft werden. Die Autorinnen und Autoren können somit eine Publikation mit dem der Veröffentlichung zugrunde liegenden Forschungsdatensatz oder -set verbinden und die beschriebenen Ergebnisse direkt anhand der Datensätze überprüfbar machen. Das Rechtekmanagement bleibt bei den Autorinnen und Autoren, nicht jeder Datensatz, der eine DOI erhält, ist zwangsläufig frei zugänglich, im Einzelfall kann eine Zustimmung des Rechteinhabers erforderlich sein. Es ist aber durch

die Speicherung im Datenrepositorium garantiert, dass der Datensatz nicht verändert werden kann.

Die DOI-Vergabe erfolgt durch die Bibliothek. Zu diesem Zweck hat die Bibliothek einen Vertrag mit der Technischen Informationsbibliothek (TIB) geschlossen, die als deutsche Agentur für die Vergabe von DOIs für Forschungsdaten fungiert. Die TIB ist Mitglied bei DataCite, einem weltweiten Zusammenschluss von Organisationen, die die Speicherung von Forschungsdaten und deren dauerhafte Verfügbarkeit standardisieren und konkret umsetzen.

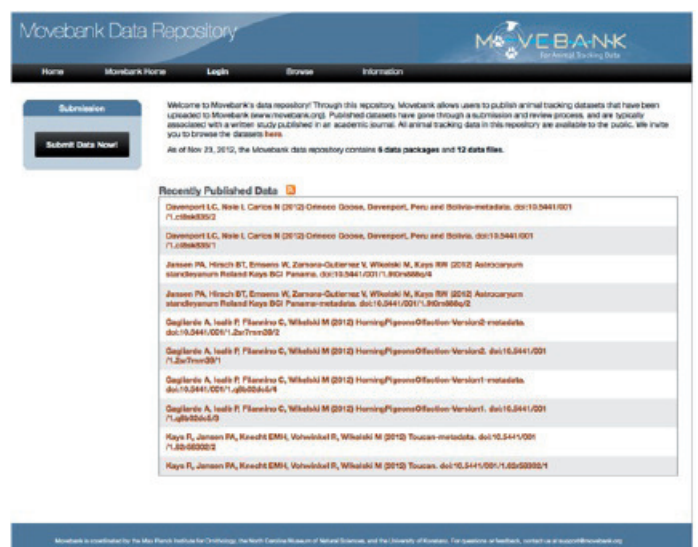


Abb. 2: Movebank Data Repository: Startseite

Studies

Data Add Sharing Upload Data Download Manage Deployments Manage Live Feeds

Study Details

| | |
|---------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Study Name | Astrocaryum standleyanum Roland Kays BCI Panama |
| Contact Person | Rolandisimo (Roland Kays) |
| Principal Investigator | Rolandisimo (Roland Kays) |
| Citation | Jansen, P.A., Hirsch, B.T., Emsens, W-J., Zamora-Gutierrez, V., Wikelski, M., Kays, R.W., 2012, Thieving rodents as substitute dispersers of megafaunal seeds: Proceedings of the National Academy of Sciences of the United States of America, v. 109, i. 31, doi:10.1073/pnas.1205184109. |
| Acknowledgements | Ben Hirsch, Patrick Janson |
| Grants used | NSF grant DEB0717071 |
| License Terms | These data have been published by the Movebank Data Repository with DOI 10.5441/001/1.9t0m888q. See www.datarepository.movebank.org/handle/10255/move.291 . |
| Study Objectives | to test the effectiveness of agoutis as dispersers of <i>Astrocaryum</i> seeds. Because <i>Astrocaryum</i> seeds are relatively large, it has been assumed that this plant species was originally dispersed by extinct Gompothers. We tested the hypothesis that medium sized rodents could fill the same ecological niche as the extinct mammals by recording complete dispersal kernels of <i>Astrocaryum</i> seeds using telemetric seed tags. We found that multi-step dispersal caused by massive agouti cache theft led to much larger dispersal distances than previously assumed. |
| Study Reference Location | |
| Longitude | -79.846 |
| Latitude | 9.159 |
| Study Statistics | |
| Number of Animals | 224 |
| Number of Tags | 224 |
| Number of Deployments | 224 |
| Number of Locations | 1649 |
| Time of First Location | 2010-05-16 06:51:00.000 |
| Time of Last Location | 2011-04-28 08:48:00.000 |
| Sensor Types | Radio Transmitter |
| Taxa | |

Die Bibliothek hat mit dem Aufbau eines Datenrepositoriums ein neues Aufgabenfeld betreten. Die Speicherung, Verfügbarmachung und langfristige Sicherung von Forschungsprimärdaten ist ein Thema, das in den vergangenen Jahren an Bedeutung gewonnen hat. Alle Forschungsfördereinrichtungen, wie z. B. auch die DFG, verlangen mittlerweile neben der Berücksichtigung der Regeln zur guten wissenschaftlichen Praxis bei Anträgen auch Aussagen zum Forschungsdatenmanagement im Projekt und in der beantragenden Institution. Die Möglichkeit, große Datenmengen schnell verarbeiten zu können, produziert neue Daten. Die beiden Ziele Nachnutzbarkeit von Daten sowie die Überprüfbarkeit von Ergebnissen stehen im Mittelpunkt der Anforderungen. Die Verknüpfung von Forschungsprimärdatensätzen mit Publikationen wird zunehmen.

Bibliotheken haben neben der Aufgabe, für eine schnelle und möglichst unkomplizierte Literatur- und Informationsversorgung zu sorgen, auch die Funktion, als Gedächtnisorganisation zu fungieren.

Dauerhafte Speicherung von Publikationen übernimmt die Bibliothek der Universität Konstanz schon beim Institutional Repository KOPS. KOPS speichert Publikationen von Angehörigen der Universität Konstanz, wo urheberrechtlich möglich und von den Autorinnen und Autoren gewünscht, dauerhaft als Volltext und stellt diesen weltweit öffentlich zur Verfügung. Eine ähnliche Aufgabe wird die Bibliothek im Kontext des Serviceverbundes Kommunikation, Information, Medien (KIM) für Forschungsprimärdaten übernehmen. Die Beschäftigung mit disziplinspezifischen Anforderungen, mit Rechtemanagement und Versionierung sowie die Bewältigung der Datenmengen werden anstehende Aufgaben sein. Karlheinz Pappenberger hat seit Beginn dieses Jahres neben seinen Aufgaben als Fachreferent die Funktion als Referent für Forschungsdaten übernommen. In der Bibliothek aktuell wird voraussichtlich noch des Öfteren über das Thema zu lesen sein.

1 http://www.allianzinitiative.de/de/handlungsfelder/virtuelle_forschungsumgebungen/ [14.4.2013]

2 www.datarepository.movebank.org