

# Potenziale der KI-basierten Biodiversitätsforschung

Von der Lagerfeuer-Idee zur Flora-Incognita-App

Im Rahmen der interdisziplinären KI-Forschungsgruppe „KI4Biodiv – Künstliche Intelligenz in der Biodiversitätsforschung“ werden am Max-Planck-Institut für Biogeochemie und an der Technischen Universität Ilmenau KI-Methoden entwickelt und verbessert, um die Biodiversität in verschiedenen Lebensräumen effizient, schnell und automatisiert überwachen und vorhersagen zu können. In der Flora-Incognita-App, mit der eine automatische Pflanzenbestimmung möglich ist, kommen diese entwickelten Technologien zur Anwendung. Zusammen mit Standort und Zeitpunkt der Pflanzenfunde wird ein großer Datensatz aufgebaut, der es ermöglicht, langfristige ökologische Veränderungen in der Verbreitung von Pflanzenarten zu dokumentieren.



”

*Ein Kollege hatte bei einer Exkursion eines Abends am Lagerfeuer eine Sternenhimmel-App dabei, mit der man Sternbilder bestimmen kann. Da kam mir die Idee, wie großartig so etwas auch für Pflanzen wäre.*

**Jana Wäldchen**, Gruppenleiterin im Projekt Flora Incognita am Max-Planck-Institut für Biogeochemie in Jena

## Worum geht es in Ihrem Projekt Flora Incognita?

**Jana Wäldchen:** In unserer Forschungsgruppe haben wir die Flora-Incognita-App entwickelt, mit der Nutzerinnen und Nutzer mithilfe einer Smartphonekamera Pflanzen bestimmen können, und die gleichzeitig dabei hilft, einen großen Datensatz zu Vorkommen und Verbreitung von Pflanzenarten zu dokumentieren. Grundvoraussetzung einer bildbasierten automatischen Pflanzenbestimmung ist ein riesiger Bilddatensatz zum Trainieren der neuronalen Netze. Für die Erstellung des Bilddatensatzes haben wir zunächst die Flora-Capture-App entwickelt, mit der es möglich ist, strukturierte Pflanzenbilder aufzunehmen. Zunächst war diese App nur für unsere Botanikerinnen und Botaniker gedacht. Sie sollten eine Möglichkeit haben, die Pflanzen in der Natur aufzunehmen. Dann haben wir aber gesehen, wie toll die App funktioniert. So ist ein richtiges Citizen-Science-Projekt daraus geworden. Nutzerinnen und Nutzer können Pflanzen aus verschiedenen Perspektiven fotografieren, die sie beispielsweise beim Spaziergang oder einer Bergtour sehen. Die Bilder werden an uns übermittelt und von den Botanikerinnen und Botanikern verifiziert. Die so erhaltenen Bilder dienen als wichtige Grundlage unserer Forschung, da wir Bilder der Pflanze aus verschiedenen und auch ungewöhnlichen Perspektiven brauchen, also nicht immer nur die Blüte, sondern vielleicht auch mal Blätter oder Bilder von verschiedenen Wachstumsphasen der Pflanze. So war es erstmals möglich zu erforschen, welche Perspektiven der Pflanze am besten geeignet sind, um die automatisierte Bilderkennung bestmöglich trainieren und realisieren zu können. Die Ergebnisse unserer Forschung finden direkte Anwendung in der

Flora-Incognita-App. Mit dieser App ist es möglich, einfach und schnell Pflanzen zu bestimmen. Oftmals genügt nur ein Bild, und die Nutzerinnen und Nutzer erhalten den Namen der Pflanze sowie weiterführende Informationen wie Merkmale, Schutzstatus und Verbreitungsgebiete. Inzwischen benutzen die App über fünf Millionen Bürgerinnen und Bürger.

### Wie ist die Idee für dieses Anwendungsprojekt entstanden?

**Wäldchen:** Als Biologin musste ich während meiner Ausbildung schon immer viele Pflanzen bestimmen. Ich konnte gut nachvollziehen, dass die herkömmliche Bestimmung mithilfe eines Bestimmungsschlüssels für viele Interessierte sehr schwer ist. Die Idee zu dieser App ist dann eines Abends bei einer Exkursion am Lagerfeuer entstanden: Ein Kollege hatte eine Sternenhimmel-App dabei, mit der man verschiedene Sternbilder am Himmel verfolgen kann. Da kam mir die Idee, wie großartig so etwas auch für Pflanzen wäre. Eine initiale Literaturrecherche im Jahr 2012 ergab, dass die automatische Pflanzenbestimmung wirklich noch in den Kinderschuhen stand. Zwar arbeiteten schon lange Wissenschaftlerinnen und Wissenschaftler daran, Pflanzen anhand von Bildern automatisch zu erkennen, aber die bis dahin entwickelten Methoden ermöglichten eine automatische Bestimmung von weniger als 100 Arten. Zudem mussten die Bilder oft auf weißem Hintergrund unter Labormethoden aufgenommen werden. Allein in Deutschland haben wir mehr als 3000 Pflanzenarten. Die bisher dazu verwendeten Methoden waren absolut noch nicht praxistauglich.

Das ursprüngliche Ziel des Flora-Incognita-Projektes war es dann, die thüringische Flora zu modellieren, um eine KI zu trainieren, mit der es möglich ist, die gesamte thüringische Flora mit mehr als 1.700 Arten automatisch zu bestimmen. In der ersten Projektphase sammelten wir dazu mehr als 500.000 Pflanzenbilder. Nun haben wir die Flora-Incognita-App auf die gesamte mitteleuropäische Flora erweitert mit mehr als einer Million Pflanzenbildern.

### Das Projekt verknüpft die Erforschung und den Transfer von KI-Methoden durch die Entwicklung einer praktischen Anwendung. Wie konnten Sie das Projekt realisieren und was sind die Voraussetzungen, damit solche KI-Projekte gelingen?

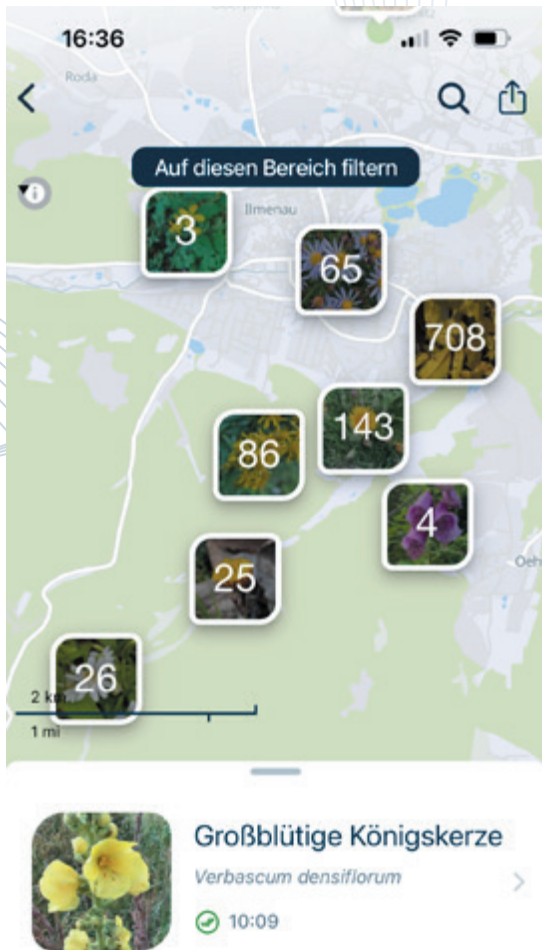
**Wäldchen:** Die Flora-Incognita-App ist im Rahmen einer Projektförderung durch das Bundesministerium für Bildung und Forschung (BMBF), das Bundesamt für Naturschutz



Die App analysiert für die genaue Bestimmung Pflanzenbilder mit einer Abfolge von tiefen neuronalen Netzen.

(BfN), das Bundesumweltministerium (BMUV) und die Stiftung Naturschutz Thüringen entstanden. Mit Professor Mäder von der Technischen Universität Ilmenau fand ich einen Forschungspartner, der der Forschungsidee zugestimmt hat. Gemeinsam habe ich mich mit ihm 2012 auf eine Fördermittelausschreibung im Zusammenhang mit der Biodiversitätsstrategie der Bundesregierung beworben. Die gemeinsame Ausschreibung zwischen Umwelt- und Forschungsministerium (ministeriumsübergreifende Ausschreibung) ermöglichte ein umsetzungsorientiertes Forschungsprojekt. Durch die Verknüpfung von Forschung und konkreter Umsetzung konnte in relativ kurzer Zeit eine praktische Anwendung entwickelt werden, die Bürgerinnen und Bürger aktiv beim Artenschutz einbezieht. Anders als bei forschungszentrierten Projekten üblich, war der Zeitrahmen von Erforschung, Entwicklung und letztendlich zur Anwendung hier sehr kurz.

Zentrale Voraussetzung, damit solche Projekte gelingen, ist die Zusammenarbeit in interdisziplinären Teams: Insgesamt besteht unser Team aktuell aus Expertinnen und Experten mit ganz unterschiedlichen Fachgebieten, wie Ökologie, Botanik, Informatik und Medieninformatik. Für die Infor-



matikerinnen und Informatiker in unseren Projekten bietet die Biodiversitätsforschung ein super Anwendungsfeld für ihre entwickelten Methoden. Gleichzeitig ist es wichtig, dass Domänenexpertinnen und -experten, wie wir Biologinnen und Biologen, sowie Botanikerinnen und Botaniker, die Potenziale der interdisziplinären Zusammenarbeit mit KI-Expertinnen und -Experten erkennen, um gemeinsam innovative Projekte umsetzen zu können. Ich denke, die Fördermittelgeber in den verschiedenen Ministerien haben das erkannt und fördern vermehrt interdisziplinäre Projekte.

### Welchen Beitrag liefern solche Projekte für die Biodiversitätsforschung?

**Wäldchen:** Der Schutz der Artenvielfalt steht im Vergleich zum Klimawandel in der öffentlichen Problemwahrnehmung oftmals zurück. Das Besondere an unserer Forschungstätigkeit und den von uns entwickelten Apps ist daher der Mitmach-Charakter. Jeder und jede kann ohne jegliches Vorwissen Pflanzen aufnehmen und bei der Weiterentwicklung der automatisierten Überwachung der biologischen Vielfalt unterstützen. Mein Hauptziel war, die Pflanzenbestimmung zu vereinfachen und den Menschen so auch einen leichteren Zugang zu ermöglichen, dass sie die Vielfalt sehen und ein Bewusstsein dafür entwickeln.

### Künstliche Intelligenz in der Biodiversitätsforschung

Im Rahmen der interdisziplinären KI-Forschungsgruppe „KI4Biodiv – Künstliche Intelligenz in der Biodiversitätsforschung“ werden am Max-Planck-Institut für Biogeochemie und an der Technischen Universität Ilmenau KI-Methoden entwickelt, um die Biodiversität in verschiedenen Lebensräumen effizient, schnell und automatisiert zu überwachen. Im Bereich des Arten- und Populationsmonitorings werden dazu Lernalgorithmen sowie Netzwerkarchitekturen speziell für Biodiversitätsdaten angewendet und weiterentwickelt. Zudem erarbeitet die Forschungsgruppe Methoden und Software, um die Anwendung von KI in der Biologie und Ökologie auszuweiten. Ein Projekt ist die Flora-Incognita-App, mit der Nutzerinnen und Nutzer mittels Smartphone Fotos von Pflanzen aufnehmen können. Anschließend wird die unbekannte Pflanze automatisch bestimmt.

Wird die Pflanze erfolgreich bestimmt, erhält man zusätzliche Informationen zu deren Merkmalen, Verbreitung oder Schutzstatus. Die Entwicklerinnen und Entwickler der App können mithilfe dieses Citizen-Science-Ansatzes, der die aktive Bürgerbeteiligung ermöglicht, sehr viel mehr Bilddateien zusammentragen, mit deren Hilfe KI-Algorithmen weiter trainiert werden. Durch die automatisierte Identifizierung von Arten können zudem Veränderungen von Pflanzenbeständen in nahezu Echtzeit überwacht und wichtige Schlussfolgerungen zu bedrohten und invasiven Arten abgeleitet werden (z.B. Verbreitungsgebiete, Verschiebung von Blütezeiträumen, mögliche Schutzmaßnahmen).

Perspektivisch können diese Daten mit anderen Datensätzen (z.B. Wetter- oder Satellitendaten) abgeglichen und Zusammenhänge zu Klimaveränderungen untersucht werden. Die entwickelten KI-Technologien für die automatische Pflanzenbestimmung können zukünftig auch für andere Artengruppen wie Insekten, Algen oder Pilze übertragen werden. Die Flora-Incognita-App ist im Rahmen einer Projektförderung durch das Bundesministerium für Bildung und Forschung (BMBF), das Bundesamt für Naturschutz (BfN), das Bundesumweltministerium (BMUV), die Stiftung Naturschutz Thüringen und das Thüringer Umweltministerium entwickelt worden.

Gleichzeitig ist es für uns sehr nützlich, dass wir zusätzlich millionenfach neue Beobachtungsdaten, also Observationspunkte, bekommen. An einem guten Tag am Wochenende im Frühling oder Sommer erhalten wir mehr als 300.000 Bestimmungsanfragen. Im Rahmen des KI4Biodiv-Verbundprojektes untersuchen wir dann, inwieweit wir mit diesen Observationspunkten Untersuchungen etwa zur Phänologie der Pflanzen durchführen können, also zum Wachstum und zu Entwicklungserscheinungen der Pflanzen im Jahresablauf. Die meisten Nutzerinnen und Nutzer der App machen Fotoaufnahmen von blühenden Pflanzen. Das sehen wir an der Anzahl der Beobachtungen pro Pflanzenart. Wann und wo welche Pflanzenarten blühen, ist u.a. abhängig von klimatischen Gegebenheiten. Wird die App

über mehrere Jahre verwendet, können wir langfristig die Beobachtungsdaten mit Klimaveränderungen abgleichen. So können wir beispielsweise Zeitverschiebungen bei Blütezeiträumen erkennen und somit Auswirkungen des Klimawandels auf biologische Systeme besser untersuchen. Aktuelle Untersuchungen zeigen, dass die Beobachtungsdaten mit der Flora-Incognita-App die phänologischen Erhebungen des Deutschen Wetterdienstes sehr gut ergänzen können.

### Wie erfolgt das Biodiversitätsmonitoring?

**Wäldchen:** Die nutzergenerierten Daten ermöglichen uns hier also, unsere Daten weiter zu ergänzen. Perspektivisch wollen wir versuchen, die nutzergenerierten Daten auch mit Satellitendaten zu verknüpfen, etwa um die Verbreitungsgebiete invasiver Arten zu überwachen und die automatisierte Biodiversitätsüberwachung insgesamt weiterzuentwickeln.

Das kontinuierliche Biodiversitätsmonitoring ist sehr wichtig, um Trends zu erkennen: Wo stehen wir gerade, wird es besser, wird es schlechter? Erst aus der Beobachtung dieser Entwicklung lassen sich konkrete Maßnahmen ableiten: etwa dann, wenn die Artenvielfalt in Gebieten abnimmt oder die Ausbreitung invasiver Arten zunimmt und dadurch andere Arten verdrängt werden. Auf Basis dieser Erkenntnisse können Bewirtschaftungsmaßnahmen festgelegt oder die Wirksamkeit von Bewirtschaftungspraktiken zur Erhaltung der biologischen Vielfalt bewertet werden.

### Was sind die nächsten Herausforderungen in der KI-basierten Biodiversitätsforschung?

**Wäldchen:** Mithilfe der trainierten KI versuchen wir nun, eine noch feinere Bestimmung von Pflanzenarten durchzuführen. Wir haben derzeit ungefähr 4.800 Arten, die wir in unserem Datensatz unterstützen. Viele Arten sind aber sehr ähnlich, da versuchen wir Algorithmen für die App zu entwickeln, um die Spezifikation noch genauer zu machen. Weiterhin ist es auch wichtig, Pflanzenarten in frühen Entwicklungsstadien zu erkennen. Dies kommt zum Beispiel in der Landwirtschaft zum Einsatz. Bisher wird nur automatisch zwischen Unkraut und Nutzpflanze unterschieden. Für eine nachhaltigere Landwirtschaft ist es jedoch wichtig, Unkrautarten zu unterscheiden und selektiert zu bekämpfen. Zum Beispiel naturschutzrelevante Arten stehen lassen und nur konkurrenzstarke Arten bekämpfen. Wir sehen aktuell, dass die KI-basierte Biodiversitätsforschung im Bereich der Pflanzen sehr gut funktioniert, auch bei größeren Tierarten funktioniert das schon gut. So können etwa



Herden bedrohter Arten in den Savannengebieten durch Bilderkennung automatisiert überwacht werden. Auch hier gibt es vielversprechende Vorteile des KI-basierten Monitorings: Der ganze Vorgang ist sehr viel ressourcenärmer und weniger störend für die Tiere, da Forscherteams nicht mehr selbst vor Ort Bilder aufnehmen müssen, sondern fest installierte Tierkameras die Bilddateien liefern. Bei kleineren Tierarten, etwa Insekten, ist das KI-basierte Monitoring aktuell noch recht schwierig, da die Tiere kleiner und daher schwieriger aufzunehmen sind und es viel mehr verschiedene Arten gibt. Aber auch hier sind die aktuellen Entwicklungen sehr vielversprechend.

### Sie verfolgen mit Ihrem Team einen sehr anwendungsbezogenen KI-Forschungsansatz. Wie könnte man KI-Nachwuchswissenschaftlerinnen und -wissenschaftler noch gezielter dabei unterstützen, praktische und anwendungsbezogene KI-Projekte in interdisziplinären Teams umzusetzen?

**Wäldchen:** Projekte wie Flora Incognita sollten neben der Grundlagenforschung auch in der Wissenschaft mehr Anerkennung finden und Wissenschaftlerinnen und Wissenschaftlern sollten langfristige Perspektiven gegeben werden. Es braucht einen langen Atem, im schnelllebigen Wissenschaftssystem solche Projekte aufzubauen und umzusetzen. Langfristig werden diese Projekte jedoch wichtige Daten liefern, die auch in der Grundlagenforschung Anwendung finden.



**Dr. Jana Wäldchen** ist Gruppenleiterin im Projekt Flora Incognita am Max-Planck-Institut für Biogeochemie in Jena. Sie verfügt über umfassende Expertise im Schnittstellenbereich der Biogeowissenschaften und der Informatik. Ihre Forschungsschwerpunkte sind die automatische Bestimmung von Arten, das Monitoring von Ökosystemen und die Auswertung von Citizen-Science-Daten für verschiedene ökologische Fragestellungen.