Tuexenia 35: 309–319. Göttingen 2015.
doi: 10.14471/2015.35.015, available online at www.tuexenia.de

Vegetweb 2.0 – Neuauflage eines Vegetationsdatenportals für Deutschland

Vegetweb 2.0 – remaking the national vegetation dataportal for Germany

Florian Jansen[1, *], Jörg Ewald[2] & Ute Jandt[3]

[1]Institut für Botanik und Landschaftsökologie, Ernst-Moritz-Arndt-Universität, Soldtmannstr. 15 17487 Greifswald, Germany, jansen@uni-greifswald.de;
[2]Fakultät Wald und Forstwirtschaft, Hochschule Weihenstephan-Triesdorf, Hans-Carl-von-Carlowitz-Platz 3, 85354 Freising, Germany, joerg.ewald@hswt.de;
[3]Institut für Biologie, Geobotanik & Botanischer Garten, Am Kirchtor 1, 06108 Halle, Germany, ute.jandt@botanik.uni-halle.de
**Korrespondierender Autor*

Zusammenfassung

Deutschland besitzt mit einer geschätzten Zahl von über 2,5 Millionen Vegetationsaufnahmen den weltweiten größten Schatz dieser wertvollen Biodiversitätsinformationen. Leider bleiben Sichtbarkeit und Zugänglichkeit dieses Erbes weit hinter den technischen Möglichkeiten und internationalen Standards zurück. Modernisierung und Ausbau der nationalen Vegetationsdatenbank vegetweb sind eine vordringliche Aufgabe, die im Rahmen eines Verbundprojektes in Angriff genommen wird. Unter http://www.vegetweb.de findet sich der Prototyp der Plattform für Vegetationsdaten, die bis Ende 2016 seine volle Funktionsfähigkeit erhalten soll.

Wir rufen alle Kolleginnen und Kollegen dazu auf, aktiv an der nationalen Vegetationsdatenbank für Deutschland mitzuarbeiten, indem sie ihre eigenen Daten zur Verfügung stellen oder sich darum bemühen, ihnen bekannte Datenbesitzer in ihrem Umfeld von einer Mitarbeit zu überzeugen.

Abstract

With more than 2.5 million vegetation plots Germany holds probably the biggest treasure of this important biodiversity information. Unfortunately, visibility and accessibility of this heritage remain far behind the technical possibilities and actual international standards. Modernization and development of the national vegetation database vegetweb is a priority which is tackled by a joint project of the biggest database maintainers in Germany. You can access a prototype of the platform at http://www.vegetweb.de. The platform for the disclose and distribution of vegetation data will reach its full operational capacity by the end of 2016.We call on all colleagues to participate actively in the National Vegetation Database for Germany by providing own data or by persuading others to do so.

Keywords: biodiversity informatics, data sharing, ecoinformatics, phytosociology

1. Einleitung

Die Verfügbarkeit großer Mengen von Vegetationsaufnahmen mit ihren vielfältigen Angaben zu Artenzusammensetzung, geografischer Lokalisierung und zu den Standortbedingungen, eröffnet Möglichkeiten der vergleichenden Analyse, nicht nur im Rahmen der klas-

Manuskript eingereicht am 23. April 2015, angenommen am 16. Juni 2015
Koordinierende Editorin: Brigitta Erschbamer

sischen Pflanzensoziologie (EWALD 2003), sondern weit darüber hinaus, z. B. im Naturschutz, in der Modellierung von Landschaftsveränderungen (JANDT et al. 2011, EWALD et al. 2013, DIEKMANN et al. 2014), der Ökologie von Neophyten (JANSEN et al. 2011b) oder der Makroökologie (BECK et al. 2012). Gerade in der Kombination mit anderen Datenbanken (z. B. Merkmalsdatenbanken von Pflanzenarten oder weiteren Beobachtungsdaten) ermöglichen sie aktuelle Fortschritte in der Biodiversitätsforschung (EWALD 2005, CURRY & HUMPHRIES 2007, BÖHNKE et al. 2014, BREITSCHWERDT et al. 2014).

Weltweit wurden von Vegetationskundlern während der letzten rund 100 Jahre viele Millionen Vegetationsaufnahmen angefertigt und zunehmend in digitalen Datenbanken verfügbar gemacht (DENGLER et al. 2011). Während in anderen Ländern wie den Niederlanden (SCHAMINÉE et al. 2012) und Tschechien (CHYTRÝ & RAFAJOVÁ 2003) ein Großteil der Vegetationsaufnahmen in nationalen Datenbanken vereint wurde, existieren in Deutschland viele Dutzend unabhängige Datenbanken mit oft nur wenigen Vegetationsaufnahmen (JANSEN et al. 2011a). Für Analysen, die über die jeweils eigenen Projekte hinausgehen, ergeben sich daraus zahlreiche Schwierigkeiten:

a) Die für die jeweilige Fragestellung interessanten Datenbanken müssen erst einmal in Erfahrung gebracht und die Kontaktpersonen ausfindig gemacht werden,
b) die Bedingungen für die Nutzung der Daten müssen individuell ausgehandelt werden,
c) das Filtern relevanter Aufnahmen muss von jedem Datengeber individuell vorgenommen werden,
d) die Anpassung der Datenformate und -inhalte muss für jedes Auswertungsprojekt von neuem beginnen. Letzteres betrifft insbesondere die Harmonisierung der Sippen-Taxonomie, die einen ganz erheblichen Arbeitsaufwand darstellt.

Bezüglich des Findens relevanter Datenbanken bedeutet der Global Index of Vegetation-Plot Databases (GIVD, http://www.givd.info, DENGLER et al. 2011) einen wichtigen Fortschritt. Er bietet mit derzeit 236 registrierten Datenbanken und 3,6 Millionen registrierten Aufnahmen einen weltweiten Überblick über vorhandene Vegetationsdatenbanken. Zurzeit sind in dieser Metadatenbank 35 deutsche Vegetationsdatenbanken registriert, die zwischen 23 und 120.000 Plots enthalten. Viele wichtige Informationen wie Geltungsbereich der Datenbank, Verteilung auf Formationen, Alter der Vegetationsaufnahmen, Verteilung der Aufnahmeflächengrößen, Datenformate und anderes sind für jeden Besucher der Webseite einsehbar. Wo Daten hilfreich erscheinen, können über die hinterlegten Kontaktdaten Anfragen zur Datennutzung getätigt werden. Für die aufgeführten Problempunkte b) bis d) ist GIVD hingegen keine Hilfe. So müssen die Bedingungen zur Datennutzung weiterhin für jede Datenbank individuell ausgehandelt werden, und die Harmonisierung von Datenformaten und verwendeter Sippen-Taxonomie jedes Mal von Neuem erfolgen. Gerade die Anpassung der Taxonomie, die über eine Synonymisierung von Namen weit hinaus geht und die Berücksichtigung unterschiedlicher taxonomischer Konzepte (BERENDSOHN 1995) und unterschiedlich fein differenzierter taxonomischer Niveaus beinhaltet, macht das Zusammenfassen von Vegetationsdatenbanken zu einem mühsamen Geschäft (JANSEN & DENGLER 2010). Hier kann eine zentrale Datenhaltung entscheidende Unterstützung bieten.

Schon früh gab es im Rahmen der Arbeitsgruppe Vegetationsdatenbanken Bemühungen zur Etablierung einer nationalen Vegetationsdatenbank für Deutschland, die zur Entwicklung von vegetweb 1.0 führten (EWALD et al. 2006, 2012). Das Archiv wurde nach dem „open data"-Prinzip der digitalen Allmende (HERB 2012) aufgebaut, die Aufnahmen unter eine „creative commons"-Lizenz gestellt, die eine unbeschränkte Nutzung für nicht-kommerzielle Zwecke erlaubt. Den Kernbestand dieser offenen Daten bilden die von Udo Bohn (†) am

Bundesamt für Naturschutz gesammelten *Luzulo-Fagion*-Aufnahmen (n = 2.353), die Datenbank bayerischer Bergwälder (EWALD 1995, n = 4.934), die von Dirk Hinterlang bereitgestellten Vegetationsaufnahmen des Nordrhein-Westfälischen Landesamtes für Natur, Umwelt und Verbraucherschutz LANUV (n = 14.856), die an der Bayerischen Landesanstalt für Wald und Forstwirtschaft erstellte Naturwaldreservats-Datenbank (ABS et al. 2008, n = 1.708) sowie die für die Synopsis der *Dicrano-Pinion*-Gesellschaften Deutschlands gesammelten Aufnahmen (HEINKEN 2008, n = 575). Mit der Floristisch-soziologischen Arbeitsgemeinschaft wurde 2005 vereinbart, dass alle in der Zeitschrift Tuexenia abgedruckten Einzelaufnahmen auf Kosten der AG in vegetweb frei verfügbar gemacht werden (EWALD 2005), was für die Bände 25–30 Sven Conrad übernahm (EWALD et al. 2007). Über diesen beständigen, aber geringen Zuwachs hinaus wurden trotz grundsätzlich bestehender technischer Möglichkeiten keine neuen Datenbanken in vegetweb eingespeist. Mit der Vegetationsdatenbank von Mecklenburg-Vorpommern (JANSEN et al. 2012) und der German Reference Vegetation Database (JANDT & BRUELHEIDE 2012) existieren sogar Datenbanken, die deutlich größere Datenmengen enthalten, allerdings nicht in frei verfügbarer Form. Auch blieb die Verwendung der insgesamt ca. 25.000 Vegetationsaufnahmen aus vegetweb 1.0 hinter den Erwartungen zurück.

Im Rahmen eines über die Nationale Strategie zur biologischen Vielfalt geförderten Projektes haben sich nun die Betreuer der drei größten deutschen Vegetationsdatenbanken zusammen getan, um der Idee einer nutzerfreundlichen nationalen Vegetationsdatenbank neuen Schwung zu verleihen. Dabei stehen die Fragen der direkten Einbeziehung der primären Datenerheber, die Möglichkeit von Datenupdates, die erleichterte Recherchierbarkeit und die Zitierfähigkeit der Primärdaten bei Auswertungen im Vordergrund.

Mit dem vorliegenden Beitrag verfolgen wir drei Ziele: (1) eine Vorstellung der neuen Plattform http://vegetweb.de, (2) die Beschreibung der Datenflüsse, Datenlizenzen und der Möglichkeit der Datenzitation, und (3) den Aufruf zur Nutzung und Erweiterung von vegetweb.de und zur Mitarbeit bei der kontinuierlichen Weiterentwicklung dieser Plattform.

2. Die Plattform

2.1 Der Vorgänger vegetweb 1.0

Neben Datenbanken wie der niederländischen (SCHAMINÉE et al. 2012, GIVD ID EU-NL-001) oder der US amerikanischen (PEET et al. 2012, GIVD ID NA-US-002) war vegetweb 1.0 (EWALD et al. 2006, 2012) bei ihrer Entstehung im Jahr 2001 eine der ersten nationalen Vegetationsdatenbanken überhaupt. Die direkte Erreichbarkeit über das Internet und die flexible Verwaltung taxonomischer Konzepte waren nur zwei Aspekte der innovativen Systemarchitektur von vegetweb 1.0. Insbesondere die seit 2005 bestehende Vereinbarung mit Tuexenia, alle Vegetationsaufnahmen, die den in Tuexenia veröffentlichten Artikeln zugrunde liegen, über vegetweb der Öffentlichkeit zur Verfügung zu stellen, stellt einen bedeutsamen Schritt zur Nachvollziehbarkeit vegetationskundlicher Analysen dar.

Dennoch blieben nach einem anfänglichen Aufbruch die Zugänge weiterer Datenbestände aus und die Nutzung stagnierte auf einem niedrigen Niveau. EWALD et al. (2012) führen technische, soziale und politische Hindernisse als Gründe an. So war der Datenimport zu umständlich, es fehlten die Ressourcen um potentielle Datengeber zu unterstützen und die Möglichkeiten für Recherche und Export von Vegetationsaufnahmen waren beschränkt. Die Euphorie der „open data“-Bewegung wich zusehends der Skepsis, dass freier Zugang lang-

fristig zu wenig Anreiz für den Datentausch bietet und sich, zumindest in einem relativ kleinen Feld wie der Vegetationskunde, ökonomisch nicht trägt. Die meisten Vegetationskundler/innen und Institutionen sind nicht bereit Daten, die man selber noch weiter pflegen und auswerten möchte, unentgeltlich der Gemeinschaft zur Verfügung zu stellen. Die föderalen Strukturen in Deutschland fördern eine Zersplitterung der Datenhaltung und der dafür aufgewendeten Mittel. Projektgebundene Fördermittel sind nur bedingt geeignet um ein Archiv langfristig zu betreiben.

2.2 Internationale Initiativen (EVA und sPlot)

Auf internationaler Ebene hat sich dagegen in den letzten Jahren Erhebliches getan. Mit dem European Vegetation Archive (EVA, http://euroveg.org/eva-database) wurde 2012 eine Initiative gegründet, die nach wenigen Jahren bereits 50 Datenbanken mit 930.000 Vegetationsaufnahmen aus 54 europäischen Ländern in einer zentralen Datenbank vereint. Die beitragenden Datenbanken bleiben dabei weiterhin bestehen und ein regelmäßiger Update-Mechanismus für neue und verbesserte Daten ist vorgesehen. Die beteiligten Kustoden der Datenbanken können zentral und formalisiert Anfragen auf Datennutzung stellen, die ein Steering Committee auf ihre Angemessenheit prüft. Danach bekommen alle Kustoden, die sich für die Datennutzungsregime „restricted“ oder „semi-restricted“ entschieden haben, eine Nachricht über die Anfrage und haben drei Wochen Zeit, der Weitergabe ihrer Daten im jeweiligen Fall zu widersprechen. Seit dem Start der EVA-Datenbank sind bereits 14 Projekte auf diese Weise mit Daten versorgt worden (http://euroveg.org/eva-database-eva-projects) und die ersten auf ihnen beruhenden Publikationen wurden eingereicht.

Weit über den europäischen Raum hinaus geht die Zielrichtung von sPlot, einer 2013 im Rahmen des German Centre for Integrative Biodiversity Research (iDiv) begründeten Initiative, die es sich zum Ziel gesetzt hat, eine globale Datenbank von Vegetationsaufnahmen aus allen Biomen zusammen zu stellen, um zum Beispiel zu analysieren, inwiefern sich die Zusammenhänge zwischen Pflanzeneigenschaften und ihrer Umwelt mit dem Klima ändern. Auch hier sind die meisten europäischen Datenbanken des EVA-Konsortiums beteiligt, außerdem wurden für sPlot bereits über 60 Vegetationsdatenbanken aus außereuropäischen Ländern zur Verfügung gestellt.

2.3 Ziele von vegetweb 2.0

Vegetweb 2.0 ist angetreten, aus den Erfahrungen dieser internationalen Projekte zu lernen und eine Struktur zu schaffen, die die vielen potentiellen Datengeber in Deutschland motiviert, ihre Daten in der einen oder anderen Weise zur Verfügung zu stellen und zugleich das Management vielfältiger Teilprojekte so einfach und transparent wie möglich zu machen.

Die Attraktivität für Datengeber soll insbesondere erhöht werden durch

- eine flexible und innovative Regelung der Nutzungsrechte
- die Möglichkeit zur dezentralen Datenpflege, bei der die Daten in vegetweb durch die Eigentümer betreut und aktualisiert werden können
- eine übersichtliche Darstellung der Inhalte und der Datenbesitzer bzw. Datenkustoden
- ein verbesserter Upload von Daten
- die Möglichkeit und Verpflichtung, die aus vegetweb genutzten Daten angemessen zu zitieren.

Die Nutzung der Plattform soll attraktiver werden insbesondere durch

- eine umfassende und flächendeckende Sammlung von Vegetationsdaten aus ganz Deutschland
- Recherchemöglichkeiten nach Lage, Vegetationstypen, Pflanzenarten, Autoren sowie Datenqualitäten (z. B. Genauigkeit der Verortung, Flächengrößenangaben oder des Erhebungszeitpunktes)
- eine stark vereinfachte Datennutzung durch Exportstandards, Abfragetools und Daten-Warenkörbe
- eine Erhöhung der Sichtbarkeit von vegetweb und der in ihr enthaltenen Datensätze

2.4 Datennutzungsvereinbarungen und Workflow

In vegetweb 1.0 existierte nur eine Datennutzungslizenz in Form eine Repositoriums gemeinfreier Daten. Alle Daten, die für vegetweb 1.0 zur Verfügung gestellt wurden können jederzeit und von jedermann uneingeschränkt genutzt werden. Während dies aus Sicht der potentiellen Datennutzer die attraktivste Datenlizenz darstellt und auch für Datengeber durchaus empfehlenswert sein kann, da ihre Daten dadurch am besten sichtbar sind und die Wahrscheinlichkeit, dass zugehörige Publikationen häufiger zitiert werden ansteigt, hat sich in der Vergangenheit gezeigt, dass es viele Gründe dafür gibt, warum potentielle Datengeber ihre Daten nicht ohne weiteres in einem Open Access Regime zur Verfügung stellen wollen.

Vegetweb 2.0 will daher die einzelnen Datengeber und ihre Datensätze in den Mittelpunkt stellen (Abb. 1) und ihnen die ganze Bandbreite möglicher Datenzugangsberechtigungen zur Auswahl anbieten:

1. Freie Daten: Die gemeinfreie Nutzung von Daten ist weiterhin eine zentrale Idee von vegetweb und soll nach der Creative Commons Lizenz CC BY-SA (http://creativecommons.org/licenses/by-sa/4.0/) geschehen, d.h. freie Nutzung für jedermann/-frau sofern die Daten angemessen gewürdigt, das heißt zitiert werden. Darüber hinaus die Möglichkeit zur Weiterverbreitung, solange dies unter den gleichen Lizenzbedingungen geschieht.

2. Geschützte Daten: Eine Weitergabe der Daten geschieht nur nach einer individuellen von der vegetweb-Plattform vermittelten schriftlichen Übereinkunft. Die potentiellen Datennutzer stellen einen formalisierten Antrag an die Datenkuratoren der gewünschten Daten. Diese können diesen Antrag innerhalb einer vorgegebenen Frist annehmen oder ablehnen. In Abbildung 2 ist das Prozedere näher erläutert. Nachdem der Nutzer Daten auf der Plattform ausgewählt hat (dazu werden Metadaten zu den Datensätzen veröffentlicht und Suchmasken angeboten), schließt er die Selektion ab und stellt eine Datennutzungsanfrage. Diese Anfrage wird dauerhaft auf der Plattform gespeichert, zusammen mit dem Nutzungstyp pro Vegetationsaufnahme, den Selektionskriterien und den mit ihrer Hilfe konkret ausgewählten Vegetationsaufnahmen. Die Anfrage enthält neben der Auswahl und der Identität des Anfragenden eine Erklärung, zu welchem Zweck die Daten genutzt werden sollen und ggf. ein Angebot an die Datengeber, etwa in Form einer Beteiligung an einer geplanten Publikation, der Generierung von Mehrwert für die genutzten Daten, oder dem Zurverfügungstellen eigener Daten im Austausch. Die Datengeber, die sich für eine geschützte Datenlizenz entschieden haben, werden vom System über die jeweilige Anfrage informiert und haben 2 bis 3 Wochen Zeit, um der Weitergabe ihrer Daten zuzustimmen, oder sie abzulehnen. Die Plattform macht keinerlei Vorgaben für diese Erklärungen, veröffentlicht aber alle Anfragen und die Antworten, so dass sich ein für alle Beteiligten optimaler Interessenausgleich im Sinne einer „Best Practice“ herausbilden kann.

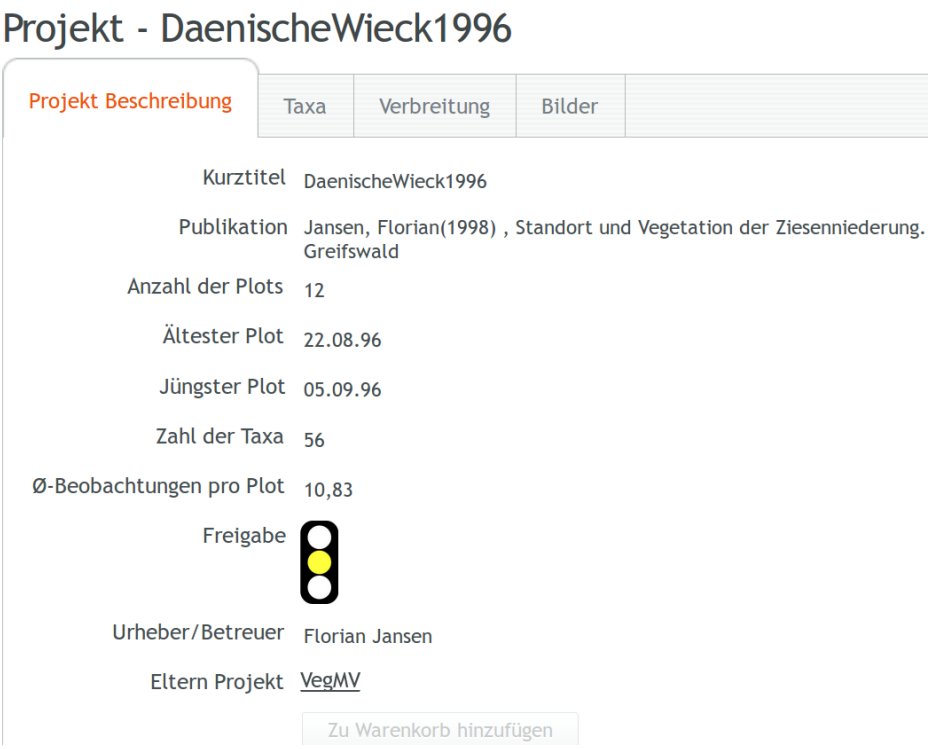


Abb. 1. Darstellung eines Einzelprojektes in vegetweb. Metadaten geben Auskunft über Inhalt und Herkunft der erfassten Daten. Kleine Auswertungen, etwa eine Liste der am häufigsten auftretenden Planzenarten und Bilder der geographischen Verbreitung der Aufnahmen und der multivariaten Ordination der Aufnahmen geben einen Überblick über den Inhalt der nicht sichtbaren Primärdaten.

Fig. 1. Representation of a primary data survey in vegetweb 2.0. Metadata and some small analyses give information about the content.

3. Gesperrte Daten: Es gibt immer wieder Gründe, warum Daten noch nicht unmittelbar zugänglich sein sollen, etwa weil man gerade selber noch dabei ist, eine erste Publikation darüber zu schreiben. Da man zu diesem Zeitpunkt jedoch den besten Überblick über seine Daten hat, möchte vegetweb auch für diese gesperrten Daten die Möglichkeit des Uploads ermöglichen, etwa um die Datenüberprüfung und -harmonisierung zu ermöglichen und die Daten für ein späteres weniger restriktives Nutzungsregime vorzubereiten. Diese Daten fließen in übergeordnete Statistiken der Plattform mit ein, sofern aus diesen nicht auf die Einzelplots zurück geschlossen werden kann. Diese Aufnahmen sind aber in keiner Weise über das System für Datennutzer freischaltbar. Das Portal behält sich vor, einen maximalen Zeitraum für diese Nicht-Lizenz einzuführen, bevor sie in ein weniger restriktives Regime überführt oder gelöscht werden müssen.

Eine aktuelle und vollständige Fassung der Datennutzungsvereinbarungen findet sich unter http://www.vegetweb.de/#!terms_view.

Mit Hilfe des sogenannten Warenkorbes (Abb. 2), in dem alle vorgenannten Informationen über die Datenanfrage gespeichert sind, kann eine individualisierte Freigabe der Daten realisiert werden. Das heißt, nur derjenige, der die Anfrage gestellt hat, kann auf die Daten

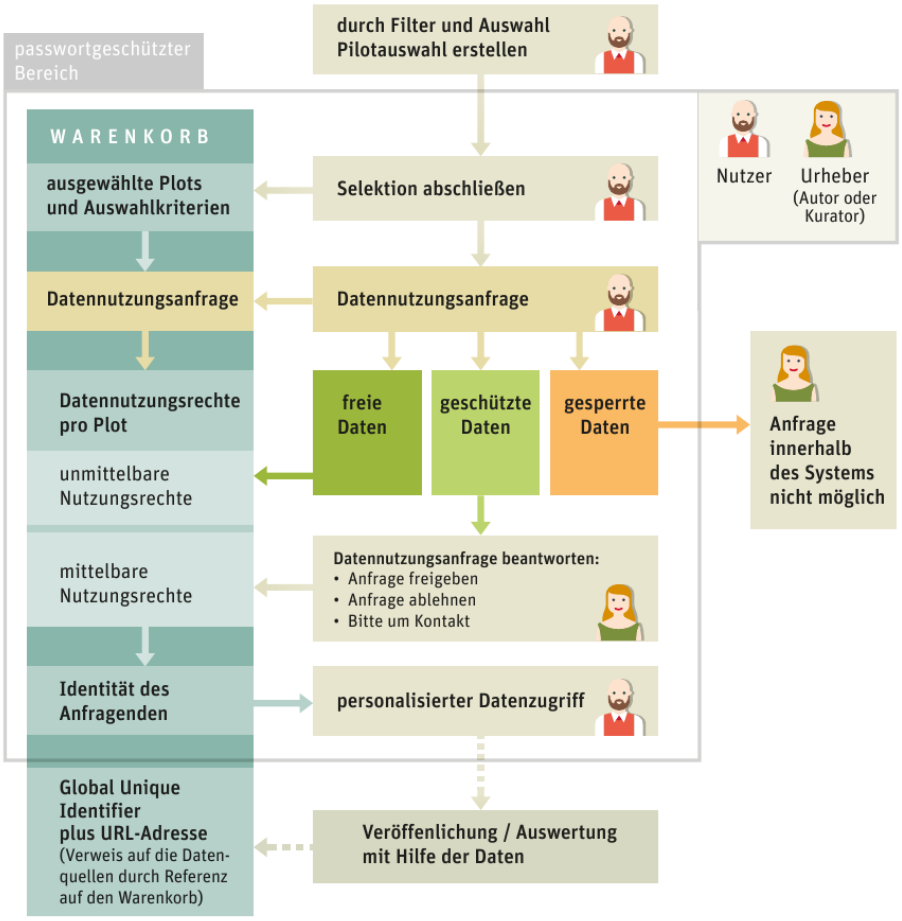


Abb. 2. Datenfluss und Umgang mit verschiedenen Datennutzungsrechten in Vegetweb 2.0. Geschützte Daten werden erst nach einer individuellen Vereinbarung zwischen Datenkustoden und Datennehmer freigegeben. Über den Warenkorb lässt sich jede Datennutzung präzise nachvollziehen und Primärdaten lassen sich mit ihrer Hilfe angemessen zitieren.

Fig. 2. Data flow and rules for data usage in vegetweb 2.0. Protected data is given free only after a negotiation between data provider and data user. Global unique identifier of data basket will guaranty the proper usage of data and offer the possibility to adequately cite the selected data.

zugreifen, dies aber jederzeit und über verschiedene Wege, sei es als flache Liste, die sich in einem Tabellenkalkulationsprogramm oder einer lokalen Datenbank öffnen lässt, als Turboveg Datenbank oder über einen sogenannten RESTful Service (representational state transfer), um direkt z. B. aus der Statistik Software R auf die Daten zugreifen zu können.

2.5 Technische Grundlagen und Datenimport

Der Plattform liegt als Datenmodell das „Toolkit“ Indicia, eine Open Source Entwicklung für die Online Erfassung Biologischer Beobachtungsdaten aus England mit einer reichen Tradition zugrunde. Im Rahmen des Projektes wurde das Datenmodell um vegetations-

spezifische Elemente wie Schichten und Deckung ergänzt. Das Indicia-Toolkit arbeitet mit der freien Datenbanksoftware PostgreSQL, die mit der Erweiterung PostGIS ein vollwertiges Geo-Informations-System bietet, so dass die Verortung von Vegetationsaufnahmen durch Punkte, Linien, Raster und Polygonflächen erfolgen kann und alle Arten von räumlichen Verschneidungen und Abfragen möglich sind. Vegetweb.de wird wie das floristische Portal Brandenburg http://www.flora-bb.de in JAVA programmiert und mit Hilfe des vaadin frameworks als dynamische Webapplikation umgesetzt.

Der Datenimport erfolgt über den XML-Export von Turboveg (HENNEKENS & SCHAMINÉE 2001), das heißt alle anderen Datenformate werden zuerst in Turboveg importiert. Auf Grundlage der Turboveg Dateien erfolgt auch die Qualitätskontrolle (Abb. 3). Als obligatorisch für jede Vegetationsaufnahme werden angesehen: ein Ort, ein Datum, eine Flächengröße und eine eindeutig referenzierte Taxonomie der aufgelisteten Pflanzen. Da bei historischen Daten nicht immer alle obligatorischen Daten vorhanden sind, wird bei einem Fehlen dieser Angaben, ein Minimum- und ein Maximumwert geschätzt bzw. beim Ort eine Unschärfe bzw. ein entsprechend unscharfes Polygon (im Notfall das Bundesland) angegeben. Die Erfassung der Arten erfolgt derzeit auf Grundlage der GermanSL (JANSEN & DENGLER 2008, http://www.botanik.uni-greifswald.de/GermanSL.html). Mittelfristig ist aber die Anbindung an die konzeptfähige taxonomische Datenbank der EDIT Plattform (http://cybertaxonomy.eu/) vorgesehen, die es ermöglichen wird, beliebige taxonomische Konzepte in die vegetweb Datenbank zu integrieren.

Fehler und Unstimmigkeiten in den Daten werden soweit wie möglich zu den Ursprungsdaten zurück getragen, um die korrigierten Daten danach erneut zu importieren. Durch die Verwendung global eindeutiger Schlüssel, sogenannter Universal Unique Identifier (uuid) und dem immer nur in eine Richtung stattfindenden Datenfluss beim Import (keine händischen Datenverbesserungen auf der Plattform) ist sichergestellt, dass nicht im Laufe der Zeit unterschiedliche Arbeitskopien der Teildatenbanken entstehen (Abb. 3).

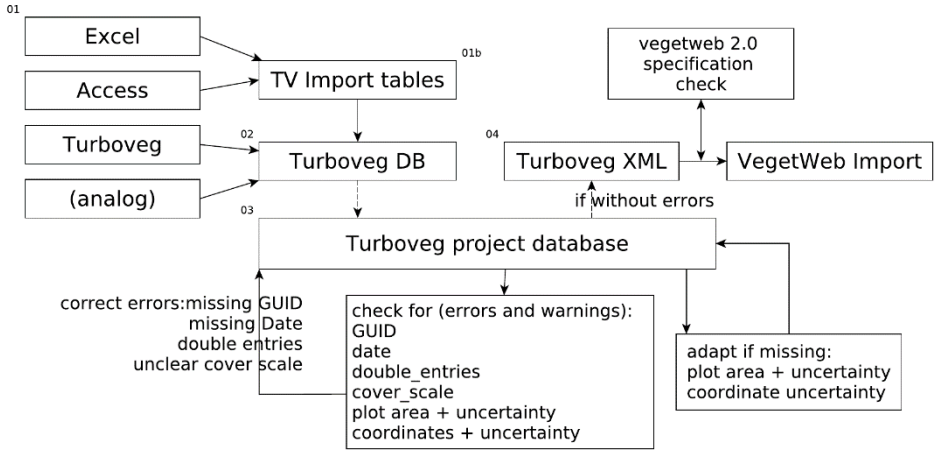
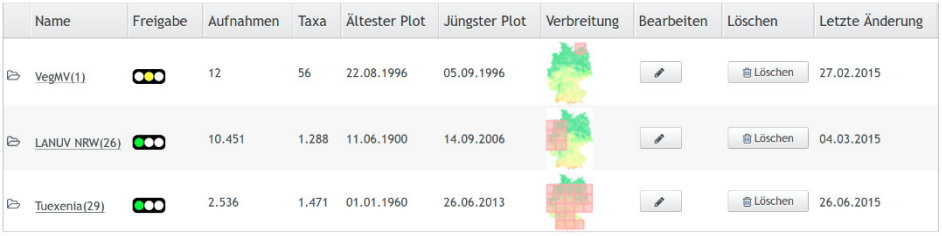


Abb. 3. Schematische Darstellung der Qualitätskontrollen beim Import der Daten in Vegetweb. Fehler und Unstimmigkeiten werden soweit wie möglich zu den Ursprungsdaten zurück getragen und die korrigierten Daten danach erneut importiert.

Fig. 3. Schematic view of data quality controls and import procedures of vegetweb 2.0. Corrections must be done in the original data to prevent branches of data developments.



	Name	Freigabe	Aufnahmen	Taxa	Ältester Plot	Jüngster Plot	Verbreitung	Bearbeiten	Löschen	Letzte Änderung
	VegMV(1)		12	56	22.08.1996	05.09.1996			Löschen	27.02.2015
	LANUV NRW(26)		10.451	1.288	11.06.1900	14.09.2006			Löschen	04.03.2015
	Tuexenia(29)		2.536	1.471	01.01.1960	26.06.2013			Löschen	26.06.2015

Abb. 4. Hierarchische Strukturierung der Primärdatenprojekte in größeren Sammlungsprojekten. Für die Weitergabe der Daten ist die jeweils restriktivste Freigabeoption bindend.

Fig. 4. Primary data surveys can be structured in hierarchical collection projects. For data clearances the most restrictive data usage rights along the hierarchy are respected.

Dennoch stellen Dubletten, das heißt die Verwendung der gleichen Vegetationsaufnahme in verschiedenen Zusammenstellungen, ein zunehmendes Problem dar. Durch die Konzentration auf die primären Datenquellen einerseits und die Möglichkeit, Primärquellen in verschiedenen, hierarchisch strukturierten Datenprojekten innerhalb der Plattform nutzen zu können, soll dem entgegen gewirkt werden (Abb. 4). Letztlich bleibt es aber in der Verantwortung des Nutzers, eine angemessene Auswahl zu treffen und Dopplungen zu vermeiden.

2.6 Zitation der Daten

Die vielleicht wichtigste Innovation von vegetweb ist die Möglichkeit, die ausgewählten Vegetationsaufnahmen zu zitieren. Jeder Warenkorb erhält eine eindeutige Identifikation und einen *Unique Resource Identifier* (*URI*, englisch für einheitlicher Bezeichner für Ressourcen), d. h. konkret eine Webadresse, über den er sich aufrufen lässt. Diese Adresse kann nun überall, wo diese speziellen Daten verwendet werden, zitiert werden. Möchte jemand mit den gleichen Daten eine weitere Analyse machen, muss er zwar (sofern es sich nicht ausschließlich um gemeinfreie Daten handelt), erneut eine Datennutzungsanfrage stellen, kann sich aber auf die alte Selektion beziehen. Die zentrale wissenschaftliche Forderung nach Wiederholbarkeit von Analysen ist damit erstmals auch für projektübergreifende Datensätze möglich.

3. Diskussion und Ausblick

CHAVAN & INGWERSEN (2009) fordern ein „data publishing framework“, um dem Mangel an primären Biodiversitätsdaten entgegenzuwirken, siehe auch BACH et al. (2011). Das Publizieren von Datensätzen (Veröffentlichung inklusive Qualitätsprüfung) und die Möglichkeit, diese Datensätze anhand einer eindeutigen ID zu zitieren, ermöglicht die Institutionalisierung eines Datenzitationsindexes analog zu den etablierten Zitationsmetriken für Manuskriptpublikationen. Mit vegetweb eröffnen wir all diese Möglichkeiten mit Hilfe der Warenkörbe und gleichzeitig die Möglichkeit, Datensätze unter unterschiedlich restriktiven und individualisierten Lizenzen zu veröffentlichen. Nach dem Ende des Projektes wird vegetweb von der Universität Greifswald an das German Centre for Integrative Biodiversity Research (iDiv) weiter gegeben. Dadurch ist eine langfristige Verfügbarkeit und Sichtbarkeit der Plattform gewährleistet. Derzeit ist es immer noch die Regel, dass von Einzelpersonen gesammelte Daten nur zu deren Lebenszeit genutzt werden können und danach verloren

sind. In Planungsbüros, Ämtern und Forschungseinrichtungen sieht es meist nicht viel besser aus. Mit vegetweb haben wir die Chance, unsere Grundlagendaten dauerhaft zu sichern und verfügbar zu machen.

Derzeit sind einige der hier erwähnten Funktionen von vegetweb noch nicht vollständig implementiert bzw. noch nicht unter Produktivbedingungen einsetzbar. Auch das Layout des Portales wird sich im Laufe des kommenden Jahres noch einmal grundlegend ändern. Wir möchten an dieser Stelle jedoch dazu aufrufen, sich schon jetzt aktiv an vegetweb 2.0 zu beteiligen, Daten beizutragen und dadurch mitzuhelfen, aus der langen vegetationskundlichen Tradition in Deutschland neue, innovative Forschungsansätze zu entwickeln, die nicht zuletzt dabei helfen können, die Biodiversität in Deutschland zu bewahren. Alle potentiellen Datengeber einschließlich der Umwelt-, Forst- und Landwirtschaftsbehörden erhalten mit vegetweb 2.0 die Gelegenheit, ihre wertvollen Biodiversitätsinformationen zugänglich zu machen, ohne die notwendige Würdigung der primären Datengeber zu vernachlässigen.

Danksagung

Wir danken Hartmut Dierschke und Thilo Heinken für die Bereitschaft, diesen Beitrag kurzfristig in die Tuexenia 35 aufzunehmen. Zwei anonymen Reviewern und der Editorin danken wir für die kritische Durchsicht.

Literatur

ABS, C., EWALD, J., WALENTOWSKI, H. & WINTER, S. (2008): Untersuchung der Schattentoleranz von Baumarten auf Grundlage der Datenbank bayerischer Naturwaldreservate. – Tuexenia 28: 23–40.

BACH, K., SCHÄFER, D., ENKE, N., SEEGER, B., GEMEINHOLZER, B. & BENDIX, J. (2011): A comparative evaluation of technical solutions for long-term data repositories in integrative biodiversity research. – Ecol. Inform. 11: 16–24.

BECK, J., BALLESTEROS-MEJIA, L., BUCHMANN, C.M., DENGLER, J., FRITZ, S.A., GRUBER, B., HOF, C., JANSEN, F., KNAPP, S., KREFT, H., SCHNEIDER, A.-K., WINTER, M. & DORMANN, C.F. (2012): What's on the horizon for macroecology? – Ecography 35: 673–683.

BERENDSOHN, W.G. (1995): The concept of potential taxa in databases. – Taxon 44: 207–212.

BÖHNKE, M., KRÖBER, W., WELK, E., WIRTH, C. & BRUELHEIDE, H. (2014): Maintenance of constant functional diversity during secondary succession of a subtropical forest in China. – J. Veg. Sci. 25: 897–911.

BREITSCHWERDT, E., JANDT, U. & BRUELHEIDE, H. (2014): Do newcomers stick to the rules of the residents? Designing trait-based community assembly tests (H. H. Bruun, Ed.). – J. Veg. Sci. 26: 219–232.

CHAVAN, V.S. & INGWERSEN, P. (2009): Towards a data publishing framework for primary biodiversity data: challenges and potentials for the biodiversity informatics community. – BMC Bioinformatics 10: S2.

CHYTRÝ, M. & RAFAJOVÁ, M. (2003): Czech National Phytosociological Database: basic statistics of the available vegetation-plot data. – Preslia 75: 1–15.

CURRY, G. & HUMPHRIES, C. (2007): Biodiversity databases: techniques, politics, and applications. - Taylor & Francis, Boca Raton, FL: 198 pp.

DENGLER, J., JANSEN, F., GLÖCKLER, F., PEET, R.K., DE CÁCERES, M., CHYTRÝ, M., EWALD, J., OLDELAND, J., LOPEZ-GONZALEZ, G., FINCKH, M., MUCINA, L., RODWELL, J.S., SCHAMINÉE, J.H.J. & SPENCER, N. (2011): The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. – J. Veg. Sci. 22: 582–597.

DIEKMANN, M., JANDT, U., ALARD, D., BLEEKER, A., CORCKET, E., GOWING, D.J.G., STEVENS, C.J. & DUPRÈ, C. (2014): Long-term changes in calcareous grassland vegetation in North-western Germany–No decline in species richness, but a shift in species composition. – Biol. Conserv. 172: 170–179.

EWALD, J. (1995): Eine vegetationskundliche Datenbank bayerischer Bergwälder. – Hoppea 56: 453–465.

EWALD, J. (2003): A critique for phytosociology. – J. Veg. Sci. 14: 291–296.

EWALD, J. (2005): Pflanzensoziologie als Beitrag zur Biodiversitätsinformatik. – Tuexenia 25: 475–483.

EWALD, J., CONRAD, S. & KLEIKAMP, M. (2007): Vegetationsaufnahmen aus Band 25 und Band 26 gehen online. – Tuexenia 27: 417–420.

EWALD, J., HENNEKENS, S.M., CONRAD, S., WOHLGEMUTH, T., JANSEN, F., JENSSEN, M., CORNELIS, J., MICHIELS, H.-G., KAYSER, J., CHYTRÝ, M., GÉGOUT, J.-C., BREUER, M., ABS, C., WALENTOWSKI, H., STARLINGER, F. & GODEFROID, S. (2013): Spatial and temporal patterns of Ellenberg values for nutrients in forests of Germany and adjacent regions - a survey based on phytosociological databases. – Tuexenia 33: 95–108.

EWALD, J., MAY, R. & KLEIKAMP, M. (2006): VegetWeb – die pflanzensoziologische online-Datenbank unter www.floraweb.de. - In: BERG C., BERGMEIER E., HÖVELMANN E., & RISTOW M. (Eds.): Ein Netzwerk für botanischen Naturschutz - neue Herausforderungen für die Botanikerinnen und Botaniker Deutschlands: 127–131. BfN- Skripten. - Bundesamt für Naturschutz, Bonn.

EWALD, J., MAY, R. & KLEIKAMP, M. (2012): VegetWeb, the national online-repository of vegetation plots from Germany. – Biodivers. Ecol. 4: 173–175.

HEINKEN, T. (2008): *Vaccinio-Piceetea* (H7) – Beerstrauch-Nadelwälder. Teil 1: *Dicrano-Pinion* – Sand- und Silikat-Kiefernwälder. – Synopsis Pflanzenges. Dtschl. 10: 1–88.

HENNEKENS, S.M. & SCHAMINÉE, J.H.J. (2001): TURBOVEG, a comprehensive data base management system for vegetation dataSoftware package for input, processing, and presentation of phytosociological data. – J. Veg. Sci. 12: 589–591.

HERB, U. (2012): Open initiatives: Offenheit in der digitalen Welt und Wissenschaft. – Universaar, Saarbrücken: 222 pp.

JANDT, U. & BRUELHEIDE, H. (2012): German Vegetation Reference Database (GVRD). – Biodivers. Ecol. 4: 355.

JANDT, U., VON WEHRDEN, H. & BRUELHEIDE, H. (2011): Exploring large vegetation databases to detect temporal trends in species occurrences. – J. Veg. Sci. 22: 957–972.

JANSEN, F. & DENGLER, J. (2008): GermanSL - eine universelle taxonomische Referenzliste für Vegetationsdatenbanken. – Tuexenia 28: 239–253.

JANSEN, F. & DENGLER, J. (2010): Plant names in vegetation databases - a neglected source of bias. – J. Veg. Sci. 21: 1179–1186.

JANSEN, F., DENGLER, J. & BERG, C. (2012): VegMV, The vegetation database of Mecklenburg-Vorpommern. – Biodivers. Ecol. 4: 149–160.

JANSEN, F., DENGLER, J., GLÖCKLER, F., CHYTRÝ, M., EWALD, J., OLDELAND, J. & SCHAMINÉE, J.H.J. (2011a): Die mitteleuropäischen Datenbanken im Global Index of vegetation-plot databases (GIVD). – Tuexenia 31: 351–366.

JANSEN, F., EWALD, J. & ZERBE, S. (2011b): Ecological preferences of alien plant species in North-Eastern Germany. – Biol. Invasions 13: 2691–2701.

PEET, R., LEE, M., JENNINGS, M. & FABER-LANGENDOEN, D. (2012): VegBank – a permanent, open-access archive for vegetation-plot data. – Biodivers. Ecol. 4: 233–241.

SCHAMINÉE, J., HENNEKENS, S.M. & OZINGA, W.A. (2012): The Dutch National Vegetation Database. – Biodivers. Ecol. 4: 201–209.

ZOBODAT - www.zobodat.at

Zoologisch-Botanische Datenbank/Zoological-Botanical Database

Digitale Literatur/Digital Literature

Zeitschrift/Journal: Tuexenia - Mitteilungen der Floristisch-soziologischen Arbeitsgemeinschaft

Jahr/Year: 2015

Band/Volume: NS_35

Autor(en)/Author(s): Jansen Florian, Ewald Jörg, Jandt Ute

Artikel/Article: Vegetweb 2.0 - Neuauflage eines Vegetationsdatenportals für Deutschland 309-319