

Die IMPETUS-Metadatenbank

Michael Süßer¹

IMPETUS – Integriertes Management-Projekt für einen Effizienten und Tragfähigen Umgang mit Süßwasser in Westafrika

Einleitung

IMPETUS ist ein interdisziplinäres und anwendungsbezogenes Forschungsvorhaben, in dem die Entwicklung von Szenarien für ein nachhaltiges Management der knappen Ressource „Wasser“ im Vordergrund steht. Das Forschungsprojekt ist in den Förderschwerpunkt „Globaler Wandel des Wasserkreislaufs (GLOWA)“ des Bundesministeriums für Bildung und Forschung (BMBF) eingebunden. Die Arbeiten werden auf der Basis eines Kompetenznetzwerkes in zwei Flusseinzugsgebieten Westafrikas durchgeführt (Abbildung 1).

An IMPETUS sind fünfzehn Institute der Universitäten Köln, Bonn, Hamburg und Rostock sowie das Max-Planck-Institut für Meteorologie in Hamburg mit insgesamt ca. 120 Wissenschaftlern beteiligt. Im ersten Projektzeitraum von 2000 bis 2003 standen überwiegend Diagnosen unterschiedlicher Aspekte des Wasserkreislaufs im Vordergrund der Betrachtungen. Darauf aufbauend werden in der zweiten dreijährigen Phase bis 2006 Prognosemethoden in bezug auf zu erwartende Veränderungen in den kommenden Jahren entwickelt. In der abschließenden letzten zweijährigen Projektphase sollen schließlich alle gesammelten Erkenntnisse in unterschiedlichen Szenarienbetrachtungen gekoppelt werden, um eine umfassende Abschätzung von Risiken und möglichen Auswirkungen globaler Veränderungen auf lokaler und regionaler Ebene zu erlauben.

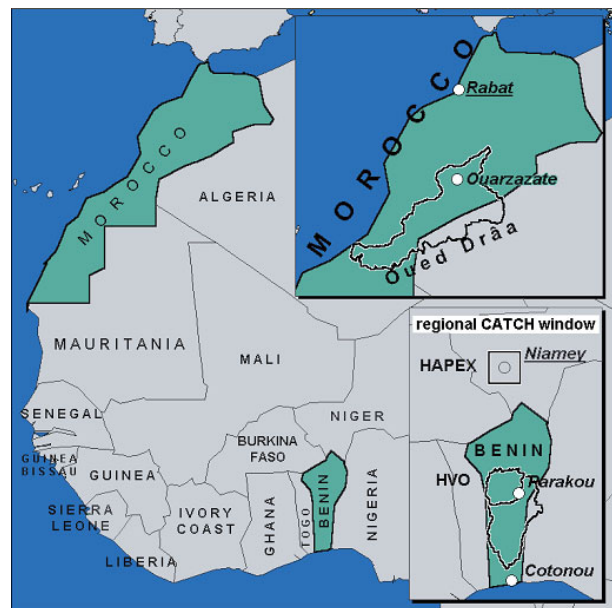


Abbildung 1: Lage der IMPETUS-Untersuchungsgebiete in Marokko und Benin.

¹ Institut für Geophysik und Meteorologie, Universität zu Köln, Kerpener Straße 13, 50923 Köln, E-Mail: suesser@meteo.uni-koeln.de

Heterogene Datensätze

Bedingt durch die interdisziplinäre Struktur von IMPETUS werden im Projekt eine Vielzahl von Datensätzen unterschiedlicher Größe, Struktur, Formatierung, usw. akquiriert und erhoben. Das Spektrum der Datensätze reicht hierbei z.B. von mehrjährigen Zeitreihen meteorologischer Messparameter über umfangreiche Satellitenbilddauswertungen auf mehreren Maßstabsebenen, vegetationskundliche Aufnahmen und Tabellen, pflanzenphysiologische Messungen auf Pflanzen- und Bestandsebene, agrarwirtschaftliche Statistiken unterschiedlicher Verwaltungsebenen, sozioökonomische Umfragen und Statistiken, ethnologische Beobachtungen, Interviews und Fallstudien bis zu umfangreichen Modellergebnissen vom globalen bis zum lokalen Maßstab. Aufgrund dieser extremen Variabilität der Datensätze ist eine zentrale Datenhaltung nicht zu verwirklichen. Um allen IMPETUS-Mitarbeitern einen Überblick und den Zugriff auf alle im Projekt erhobenen Daten zu ermöglichen wurde eine zentrale Metadatenbank entworfen, in der sämtliche notwendigen Informationen über die Datensätze verwaltet werden.

Metadatenbank

Organisation

Aus organisatorischen Gründen erfolgt die Datenhaltung in IMPETUS dezentral in den einzelnen beteiligten Instituten. Die Metadatenbank liegt zentral auf Servern der Universität zu Köln. Alle Metainformationen werden in einer Access-Datenbank erfasst, in der auch weitere Projekt bezogene Informationen gehalten werden (Mitarbeiter, Institute, Veranstaltungen, usw.). Diese Datenbank enthält außerdem Struktur und Inhalte der IMPETUS-Internetseiten. Die Websynchronisation wurde über Skripte realisiert, welche die Inhalte der Access-Datenbank in eine MySQL-Datenbank auf dem Internetserver kopieren. Die Abfrage und der Zugang zu den einzelnen Datensätzen erfolgt über das Internet.

Metadatenansätze

Bedingt durch die Vielzahl unterschiedlicher Datensätze, darunter auch nicht-räumliche Daten, wurden vorhandene Metadatenstandards, die sich i.d.R. auf Geodaten beschränken, nur als Grundlage herangezogen und an die Bedürfnisse von IMPETUS angepasst. Zur Verwendung kam vorrangig der FGDC-Standard (FGDC-STD-001-1998), der um Datenfelder für nicht räumliche Daten ergänzt wurde.

Da die Erfassung der Metadaten Aufgabe der einzelnen Wissenschaftler ist, müssen alle Hürden bei der Eingabe der Metainformationen so niedrig wie möglich gehalten werden. Daher wurde die Zahl der Datenfelder und die Zahl der Pflichtfelder auf ein Minimum reduziert. Folgende Informationen werden in der Metadatenbank erfasst:

- Metadaten – Identifizierung (Titel, Projektzuordnung, Ansprechpartner, ...)
- Metadaten – Inhalt (Beschreibung, Parameter, Erhebungszeitpunkt, Deskriptoren, ...)

- Metadaten – Datenqualität (Methoden, Prozessierung, Datenqualität, ...)
- Metadaten – Räumliche Referenz (Ort, Koordinaten, Auflösung, Maßstab, ...)
- Metadaten – Transfer (Dateiformat, Zugang, Nutzungsbeschränkungen, ...)
- Metadaten – Referenz Informationen (Erfassungsdatum, ...).

Web-Anbindung

Die Datenhaltung erfolgt in einer Access-Datenbank, auf die innerhalb der IMPETUS-Geschäftsstelle ein Netzwerkzugriff besteht.

Die einzelnen Wissenschaftler geben Metainformationen über ihre Datensätze über ein Online-Formular ein (Abbildung 2). Die Kontrolle der Eingaben (Pflichtfelder) und die Weiterleitung per Email an das Datenmanagement erfolgt mittels PHP-Skripten. In der Geschäftsstelle werden die Daten dann von Hand in die

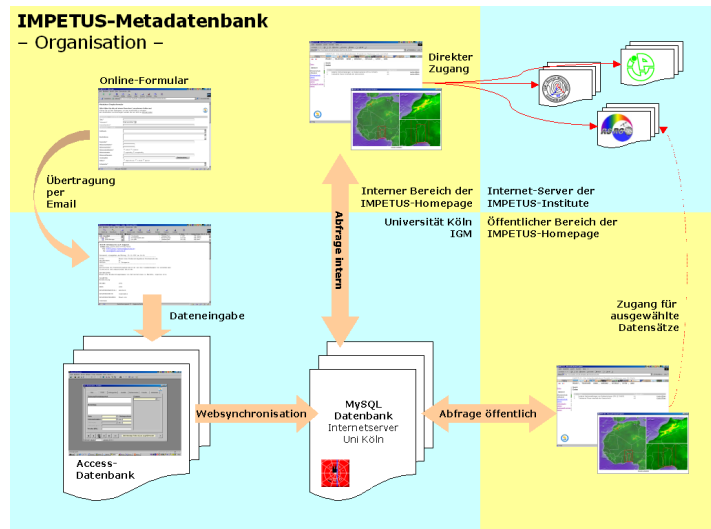


Abbildung 2: Organisationsstruktur der IMPETUS-Metadatenbank

Access-Datenbank eingegeben. Die Möglichkeit einer umfassenden Qualitätskontrolle an dieser Stelle rechtfertigt den hohen Aufwand der manuellen Dateneingabe.

Anschließend werden die Inhalte der Access-Datenbank in eine MySQL-Datenbank auf dem Internet-Server der Universität zu Köln kopiert. Die Abfrage der Metadatenbank erfolgt über PHP basierte Skripte über das Internet und ist grundsätzlich öffentlich zugänglich. Der direkte Zugriff auf die Daten selbst, die auf den unterschiedlichen Servern der beteiligten Institute gehalten werden, wird für jeden Datensatz einzeln geregelt (öffentlich, intern, kein Zugang). Ziel ist es, nach wenigen Jahren alle relevanten Daten, die in IMPETUS erhoben wurden, öffentlich zugänglich zu machen.

Ausblick

Die Struktur der IMPETUS-Metadatenbank ist entwickelt und wird derzeit von den beteiligten Wissenschaftlern getestet. Anschließend beginnt die systematische Erfassung der Metadaten. Dabei müssen auch alle Daten der ersten, dreijährigen Projektphase rückwirkend eingegeben werden.

Die Metadatenbank wird in Kürze öffentlich zugänglich sein. Der direkte Zugriff auf die einzelnen Datensätze ist aber zunächst nur intern möglich. Nach und nach werden diese aber für die Öffentlichkeit freigegeben. Außenstehende Interessenten müssen sich identifizieren und die Nutzungsbedingungen anerkennen.