OpenHyPE: OpenHygrisC-Datenverarbeitung für die Lehre

- **NRW-Förderprojekt**
- Gefördert durch das Ministerium für Landwirtschaft und Verbraucherschutz NRW (MLV)
- Laufzeit: 15.12.2021 - 30.06.2023

ENGLISH VERSION

0. Zusammenfassung

Das **Landesamt für Natur, Umwelt und Verbraucherschutz (LANUV)** des Landes **Nordrhein-Westfalen (NRW)** stellt umfangreiche quantitative und qualitative **Grundwassermessdaten** zur Verfügung. Damit wird ein Beitrag zur Erfüllung der **Europäischen Wasserrahmenrichtlinie** sowie der **EU INSPIRE-Richtlinie** für eine offene interoperable Geodateninfrastruktur geleistet.

NRW betreibt ein eigenes wasserbezogenes Datenportal **ELWAS-WEB**, das auch die Grundwasserdatenbank **HygrisC** bereitstellt. ELWAS und HygrisC sind nicht einfach zu bedienen und bieten der Öffentlichkeit nur begrenzte Möglichkeiten zur explorativen Datenanalyse.

NRW veröffentlicht jedoch einen Großteil seiner Grundwasserdaten als **offenes Datenarchiv** namens **OpenHygrisC**, das mehrere Datentabellen im **csv-Format** enthält. Insbesondere die Messdatentabelle ist **“Big Data”**, denn sie enthält alle Zeitreihen mit **mehr als 3,6 Millionen Einzelmessungen** (Tabellenzeilen). Diese große Tabelle sowie die Tabelle mit den räumlichen Koordinaten der Grundwassermessstellen erfordern den Einsatz eines **räumlichen objektrelationalen Datenbankmanagementsystems (Spatial ODBRMS)** und ein umfangreiches **Data Engineering** vor dem Einfügen in die Datenbank.

Das Ziel des **OpenHyPE-Projekts** ist es, eine erste Reihe von **Open Educational Resources (OER)** (freie Lehrmaterialien) zu entwickeln, um den Aufbau, die Befüllung und die Nutzung einer **räumlich-zeitlichen Datenbank** mit den OpenHygrisC-Daten zu trainieren. Aufgrund des **abgestuften Schwierigkeitsgrades** richtet sich das Projekt an **Schülerinnen und Schüler** von **weiterführenden Schulen** sowie **Hochschulen** im Land Nordrhein-Westfalen und darüber hinaus.

Alle verwendeten Softwareprodukte sind **Free and Open Source Software (FOSS)**. Die Datenbank, die wir **OpenHyPE DB** nennen, basiert auf **PostgreSQL / PostGIS** und bildet das Zentrum des Systems zur Analyse und Präsentation von Umweltdaten. Die Lehrmaterialien demonstrieren, wie das geographische Informationssystem **QGIS** sowie **Python-Programme** in der **JupyterLab-Entwicklungsumgebung** mit der OpenHyPE DB zusammenarbeiten, um die Daten in Form von zeitabhängigen Karten oder Zeitreihen auszuwählen, zu analysieren und darzustellen. Wir verwenden Python und Jupyter aus der **Anaconda Distribution**.

Die Anschubfinanzierung für das OpenHyPE-Projekt dient dazu, die **wertvolle Umweltdatensammlung des Landes NRW** bei jungen Menschen bekannt zu machen sowie einen Beitrag zur **interdisziplinären MINT-Förderung** im Allgemeinen und zur **Bildung für nachhaltige Entwicklung (BNE)** im Besonderen durch die **Verknüpfung von Umweltwissenschaften und**

Informatik zu leisten.

1. Einleitung

1.1 Problembeschreibung

Das Land Nordrhein-Westfalen (NRW) betreibt über das LANUV umfassende und professionelle Messnetze zur Erfassung von Umweltdaten. Im Rahmen von Open.NRW und angetrieben durch die INSPIRE-Direktive der Europäischen Union sowie weitere Direktiven wie zum Beispiel die EU-Wasserrahmenrichtlinie (WRRL) werden vom Land NRW umfangreiche Daten-Produkte offen zugänglich und frei nutzbar auf verschiedenen Plattformen zur Verfügung gestellt (Free and Open Data).

Das Land NRW ist in Deutschland ein Vorreiter bei der Bereitstellung von offenen und (kosten)freien Geodaten. Diese Daten sind ein wahrer Schatz und bilden die Grundlage für potentiell massiven Erkenntnisgewinn im Bereich Umwelt- und Naturschutz. Trotzdem scheint es so zu sein, dass nur ein vergleichsweise kleiner Personenkreis dieses Potential wirklich nutzt. Deshalb hat sich das Projekt OpenHyPE zur Aufgabe gemacht, diesen Datenbestand in die Hochschullehre einzubauen und entsprechendes frei zugängliches Lehrmaterial zu entwickeln, dass nicht nur von Studierenden sondern auch zum Teil von Schülerinnen und Schülern genutzt werden kann, um die Grundzüge der Umweltdatenverarbeitung zu lernen. Die Anschubfinanzierung soll genutzt werden, um erste Schritte der Entwicklung solchen Training-Materials umzusetzen.

Wir verfolgen dabei das Paradigma des „Problem based learning“: Die notwendigen Kenntnisse und Fähigkeiten werden anhand einer konkreten gesellschaftlich relevanten Problemstellung identifiziert und vermittelt. Die Lösung der als bedeutsam erkannten Fragestellung ist die Motivation für das Lernen.

Am Anfang wollen wir das Material anhand des Problemfelds „Grundwasserschutz“ entwickeln. Das Ministerium für Umwelt, Landwirtschaft, Natur- und Verbraucherschutz NRW (MULNV) betreibt über den „Landesbetrieb Information und Technik Nordrhein-Westfalen“ (IT.NRW) ein eigenes wasserbezogenes Datenportal namens ELWAS-WEB. Darin werden auch Daten der landesweiten Grundwasserdatenbank HygrisC vorgehalten. ELWAS und HygrisC bieten Außenstehenden nur eingeschränkte Möglichkeiten der explorativen Datenanalyse. Aus Sicht des Usability Engineering, das sich mit der Anwenderfreundlichkeit technischer Systeme befasst, sind hinsichtlich der Benutzbarkeit sowie der Datenanalyse-Möglichkeiten Verbesserungen wünschenswert, denn gerade die explorative Datenanalyse und das Data Mining helfen, Strukturen und Zusammenhänge zwischen den Daten zu erkennen. ELWAS und HygrisC sind deshalb nur bedingt geeignet, Grundlagen der Umweltdatenanalyse zu vermitteln, können aber im Unterricht als Begleitmaterial einfließen.

Auf dem Portal OpenGeodata.NRW werden umfangreiche Daten mit Raumbezug – auch Geodaten genannt – zur Verfügung gestellt, die oft einen Zeitbezug haben, wie z.B. Landnutzungsänderungen oder Messdatenreihen zur Wasserqualität. Dort liegen auch Auszüge der HygrisC-Grundwasser-Datenbank des Landes NRW, die unter dem Namen OpenHygrisC veröffentlicht werden. Diese Grundwasserdaten können in idealer Weise als Grundlage zum Aufbau einer eigenen Umweltdatenbank dienen, anhand derer die Lernenden Konzepte des Datenmanagements und der Datenanalyse kennenlernen.

1.2 Projektziele

Folgende Komponenten sollen realisiert werden:

- Entwicklung der OpenHyPE Geodatenbank auf Basis von PostgreSQL/PostGIS zur Verwaltung raum- und zeitbezogener Daten zu Grundwasserqualität und -menge
- Problembezogenes freies Online-Kursmaterial (OER), Tutorials, Video-Tutorials, Anleitungen, Programm-Code, unter Verwendung von Free and Open Source Software (FOSS):
    - Vorstellung des Landesamts für Natur, Umwelt- und Verbraucherschutz (LANUV)
    - Einführung in den Grundwasserschutz
    - Einführung in das Geographische Informationssystem QGIS
    - Einführung in die relationale Datenbank PostgreSQL und die Abfragesprache SQL
    - Einführung in die Geodatenbank-Erweiterung PostGIS
    - Einführung in die Verarbeitung von Geodaten mit der Programmiersprache Python
    - Installation des OpenHyPE Datenbank-Managementsystems
    - Diskussion des Datenmodells und Hochladen der OpenHygrisC-Daten des LANUV
    - Automatisches Erstellen von Diagrammen zu Zeitreihen der Wasserqualität
    - Automatisches Erstellen von Karten zur Grundwasserchemie
    - Erstellen einfacher Dashboards mit interaktiven Online-Grafiken und -Karten
    - Einführung in Data Mining (Deskriptive Statistik, Suchen nach Zusammenhängen)

From:
https://wiki.eolab.de/ - **HSRW EOLab Wiki**

Permanent link:
https://wiki.eolab.de/doku.php?id=eolab:openhype:german:sec_1&rev=1674308952

Last update: **2023/01/21 14:49**