

# HENRY

Hydraulic Engineering Repository

Ein Service der Bundesanstalt für Wasserbau

---

Conference Paper, Author's Postprint

**Stütz, Dominik; Kunz, Eberhard**

## **GeoValML – Ein Datenbankschema für nahezu beliebige geotechnische Kennwerte**

---

Verfügbar unter/Available at: <https://hdl.handle.net/20.500.11970/110451>

Vorgeschlagene Zitierweise/Suggested citation:

Stütz, Dominik; Kunz, Eberhard (2021): GeoValML – Ein Datenbankschema für nahezu beliebige geotechnische Kennwerte. In: Krieger, Jürgen (Hg.): 1. Fachkongress Digitale Transformation im Lebenszyklus der Verkehrsinfrastruktur. Tagungshandbuch 2021. Tübingen: expert verlag GmbH. S. 133-136.

### **Standardnutzungsbedingungen/Terms of Use:**

Die Dokumente in HENRY stehen unter der Creative Commons Lizenz CC BY 4.0, sofern keine abweichenden Nutzungsbedingungen getroffen wurden. Damit ist sowohl die kommerzielle Nutzung als auch das Teilen, die Weiterbearbeitung und Speicherung erlaubt. Das Verwenden und das Bearbeiten stehen unter der Bedingung der Namensnennung. Im Einzelfall kann eine restriktivere Lizenz gelten; dann gelten abweichend von den obigen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Documents in HENRY are made available under the Creative Commons License CC BY 4.0, if no other license is applicable. Under CC BY 4.0 commercial use and sharing, remixing, transforming, and building upon the material of the work is permitted. In some cases a different, more restrictive license may apply; if applicable the terms of the restrictive license will be binding.



**Erstveröffentlichung in Krieger, Jürgen (Hg.) (2021):  
1. Fachkongress Digitale Transformation im Lebenszyklus der Verkehrsinfrastruktur.  
Tagungshandbuch 2021. Tübingen: expert Verlag GmbH, S. 133-136.**

# GeoValML – Ein Datenbankschema für nahezu beliebige geotechnische Kennwerte

Dominik Stütz<sup>a</sup>, Eberhard Kunz<sup>a</sup>

<sup>a</sup> Bundesanstalt für Wasserbau, Karlsruhe, Deutschland

**Zusammenfassung:** Geotechnischen Versuchsdaten und Ergebnisse werden zur Beurteilung des Baugrundes u.a. bei Infrastrukturprojekten benötigt. Häufig ist dieser Datenbestand jedoch dispers verteilt oder gar nicht digital verfügbar. Der Austausch von geotechnischen Versuchsdaten und Ergebnisse ist aktuell somit schwierig. Im Rahmen der vom Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) geförderten Forschungsinitiative mFUND ([www.mfund.de](http://www.mfund.de)) hat die Bundesanstalt für Wasserbau (BAW) einen geotechnischen Begriffskatalog und das Datenbankschema GeoValML entwickelt. GeoValML steht für Geotechnical Value Markup Language. Mit der dazugehörigen IT-Infrastruktur steht nun eine Technik bereit, die einen interoperablen Austausch geotechnischer Kennwerte ermöglicht. Dies erhöht Sicherheit und Wirtschaftlichkeit von Bauprojekten. GeoValML ist der erste umfassende Begriffskatalog für eine geotechnische Kennwertdatenbank. Bereits jetzt enthält dieser über 3300 Begriffe und wird weiter stetig ausgebaut. Der Vortrag berichtet über die nutzerreife Anwenderplattform und das Entwicklungskonzept von GeoValML.

## 1 Motivation

### 1.1 Herausforderungen

Die Kenntnis über die geotechnischen Eigenschaften des Baugrunds ist für jedes Bauvorhaben zwingend erforderlich. Die Ermittlung der relevanten Kennwerte erfolgt in der Regel aufwändig und kostenintensiv mittels direkter und indirekter Aufschlüsse und Laboruntersuchungen. Die Untersuchungsergebnisse fließen in gutachterliche Beurteilungen ein und sind Grundlage der Baugrundsichtenmodelle. Durch die Bauvorhaben entstehen dabei lokal wertvolle Datensammlungen die aber in der Regel nicht öffentlich zugänglich und recherchierbar sind. Für eine deutschlandweite Sammlung und Bereitstellung der Daten fehlt zum einen ein interoperables Austauschformat und zum anderen die digitale Infrastruktur.

Bestehende Datensammlungen sind technisch heterogen aufgebaut, und verwenden viele unterschiedliche Formate und Begriffsdefinitionen, was einen Datenaustausch deutlich erschwert, wenn nicht sogar verhindert.

Die bei den Untersuchungen gewonnen Daten sind nicht nur für Bauprojekte wichtig, sondern auch für die Bereiche Ressourcen, Risikoanalyse und Umwelt.

Mit Verabschiedung des Geologiedatengesetzes (GeolDG) im Juni 2020 besteht die gesetzliche Verpflichtung, die unter anderem bei Baugrunduntersuchungen gewonnenen Daten den geologischen Landesämtern für die Veröffentlichung bereitzustellen [1]. Des Weiteren hat die Bundesanstalt für Wasserbau (BAW) als Bundesoberbehörde den Verpflichtungen dem E-Government-Gesetz (EGovG) nachzukommen. Dies schließt auch die öffentliche Bereitstellung der Baugrunddaten ein [2].

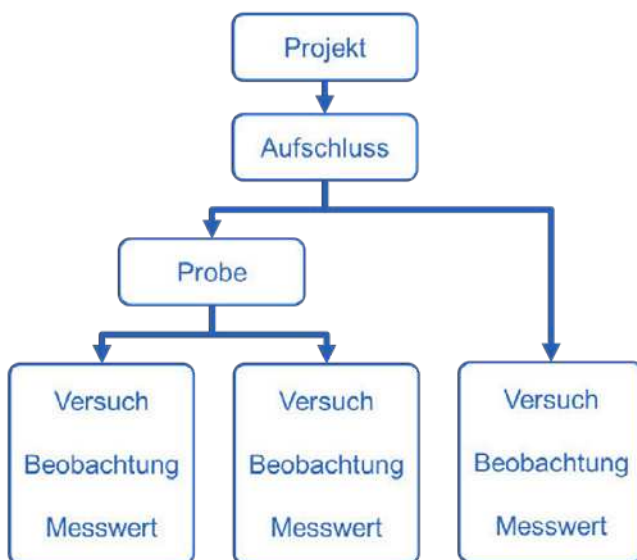
## 1.2 Zielsetzung

Aus diesen Herausforderungen lassen sich folgende Kernziele abgeleitet:

Die Daten sind so zu dokumentieren das der Nachweis ihre Herkunft auch nach Jahrzehnten möglich ist. Die verwendeten Begriffe zur Beschreibung und Einordnung der Daten sind festzulegen zu dokumentieren und zu veröffentlichen. Die Schnittstelle für den technischen Austausch der Daten sollte weltweite Konformität und Akzeptanz haben. Die verwendeten Protokolle sollten etabliert sein. Die Fachliche Schnittstelle ist so zu konzipieren, dass der Austausch zwischen Datenanbietern und Nutzern standardisiert ist.

Diese Kernziele lassen sich mit der konsequenten Umsetzung des FAIR-Prinzips erreichen. FAIR steht für Findable, Accessible, Interoperable und Reusable, zu Deutsch: Auffindbar, Zugänglich, Interoperabel und Wiederverwertbar. Das Prinzip wurde 2016 in der Fachzeitschrift Scientific Data veröffentlicht [3].

Schema geotechnische Datenbank  
Konventionell



Flexibel

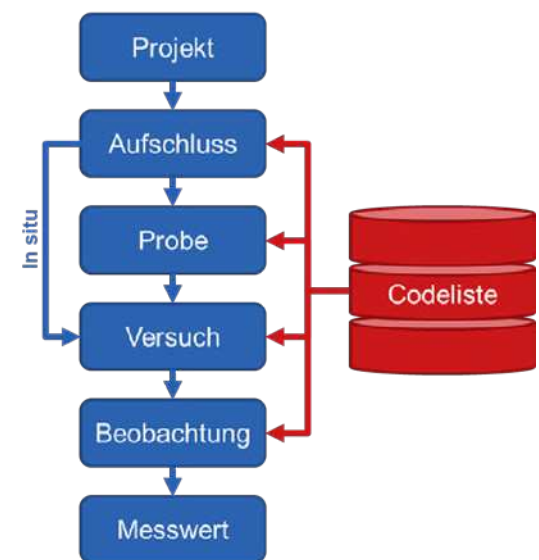


Abbildung 1: Konventionelles und flexibles Datenbankschema für eine geotechnische Datenbank

## 2 Datenbankschema

Das Datenbankschema ist eine sehr wichtige Komponente im Aufbau einer Datenbank. Einmal festgelegt kann es nach Inbetriebnahme der Datenbank nur mit hohem Aufwand verändert werden. Darum sollte dies im Vorfeld gut überlegt sein. Im Folgenden werden zwei mögliche Schemavarianten für eine geotechnische Datenbank dargelegt, welche in Abbildung 1 zu sehen sind.

### 2.1 Konventionell

Für geotechnische Kennwerte ergibt sich nach einem konventionellen Ansatz folgender möglicher Aufbau:

Jedem Projekt werden Aufschlüsse zugewiesen. Den Aufschlüssen wiederum werden Proben zugeteilt. Für alle Proben oder In situ Untersuchungen werden Tabellen für unterschiedliche Versuche angelegt. Die Versuchstabellen sind unterschiedlich aufgebaut, je nachdem welche Werte bei einem Versuch gemessen werden. In die Versuchstabellen werden schließlich die Beobachtungen und Messwerte eingetragen.

Dieser Aufbau erscheint nach der Vorgehensweise im Labor sehr naheliegend und wurde auch bei der BAW so umgesetzt. Es gibt aber einen entscheidenden Nachteil: Neue Versuche können nur ergänzt werden, indem eine neue Tabelle angelegt wird. Somit muss das Schema für neue Versuche angepasst werden. Aus diesen Erfahrungen und der in 1.2 dargelegten Zielsetzung wurde Rahmen der Forschungsinitiative mFUND das BMVI ein neues, flexibleres Schema erarbeitet.

### 2.2 Flexibel

Das flexiblere Schema ist so aufgebaut, dass auch bei neuen Versuchen keine zusätzlichen Tabellen angelegt werden müssen. Dies wird erreicht indem nicht mehr für einen Versuch Beobachtungen und Messwerte direkt eingetragen werden, sondern dies in unterschiedliche Tabellen aufgeteilt wird. Somit entsteht für jeden Messwert in der Beobachtungstabelle ein neues Tupel. Dadurch erhöht sich zwar die Datenmenge, da bestimmte Metadaten häufiger eingetragen werden müssen. Dies ist aber bei Textdaten vernachlässigbar.

Um den Austausch zu vereinfachen ist es essentiell notwendig einheitliche Begriffe (Metadaten) zu verwenden. Dafür wird ein weiteres Datenbankschema für die Metadaten benötigt. Dort werden alle Schlüssel, also definierte Begriffe, verwaltet, die für bestimmte Attribute (Eigenschaften) verwendet werden dürfen. Ist ein Schlüssel nicht vorhanden, kann dieser nicht geschrieben werden – wenn die Schlüssel-Tabelle nicht ergänzt wird. Die Ergänzung der Schlüssel Listen ist jederzeit möglich. Soll zu einer bestehenden Datenbank eine weitere Versuchsmethode hinzugefügt werden, so genügt es die entsprechenden Schlüsseleinträge zu ergänzen. In der konventionellen Labordatenbank muss hingegen hierfür eine neue Tabelle angelegt werden bzw. ist ein Eingriff in die Datenbank notwendig.

Diese Vorgehensweise sichert langfristig die Datenqualität. Vor dem Ersteinsatz einer solchen Datenbank sind die notwendigen Schlüssellisten anzulegen. Seitens der BAW wird der für den Fachbereich Geotechnik erforderliche Begriffskatalog bereitgestellt und gepflegt.

### **3 Neuentwicklung – GeoValML**

Die Geotechnical Value Markup Language ist der erste umfassende Begriffskatalog für eine geotechnische Kennwertdatenbank. Er basiert streng auf den verfügbaren Normen und Regelwerken (DIN, DGGT-Empfehlungen, ISO-Standards, ...). Bereits jetzt enthält dieser über 3300 Begriffe und wird weiter stetig ausgebaut. Damit ist der Begriffskatalog die Basis für ein interoperables Austauschformat für geotechnische Kennwerte. Er kann sehr einfach auch auf der Geotechnik fachlich nahe Bereiche ausgedehnt bzw. angepasst werden. Der Begriffskatalog ist frei verfügbar und wird bilingual in Deutsch und Englisch zur Verfügung gestellt. Das Datenbankschema wurde in Anlehnung an das Schema „Observations and Measurements“ des Open Geospatial Consortium (OGC), der internationalen Organisation zur Standardisierung von räumlichen Daten, entwickelt. Es wurde so aufgebaut, dass die Hierarchie bis zur Einzelmessung unterteilt werden kann. Alle enthaltenen Daten haben einen Raumbezug. Dies ist vor allem bei der Nutzung der Daten in BIM-Systemen von großem Vorteil.

## **4 Praktische Umsetzung**

### **4.1 Datenbank**

Die praktische Umsetzung von GeoValML ist die Geotechnical Value Database. Diese verwendet als Datenbankmanagementsystem PostgreSQL mit der Erweiterung PostGIS. PostgreSQL zählt weltweit zu den beliebtesten Datenbankmanagementsystemen und ist wie PostGIS Open-Source. Mit Hilfe der PostGIS-Erweiterung werden Messdaten als Geometrie wie z.B. der Bohrungsverlauf als Kurve im Raum abgelegt. Darüber hinaus sind räumliche Abfragen möglich.

### **4.2 Schnittstellen**

Die zentrale Schnittstelle zur GeoVal-Datenbank ist die OGC API - Features. API ist die Abkürzung für Application Programming Interface. Die API ersetzt den bisherigen Standard WFS (Web Feature Service). Die Vorteile der API sind, die Vereinfachung des Zugriffs auf verteilte Geodaten und die einfachere Integrierbarkeit in beliebige Webanwendungen und Prozesse. Die GeoVal Datenbank erhält somit die modernste OGC-konforme Schnittstelle. Die Webseite der Schnittstelle ist unter folgender Webadresse erreichbar: [www.baugrund-daten.baw.de](http://www.baugrund-daten.baw.de)

Um Daten auch manuell in die Datenbank einpflegen zu können, wird das Geotechnical Value Data Service Tool (GeoVal-DST) entwickelt. Die Besonderheit des Geo-Val-DST ist, dass sich die Eingabemasken für die Dateneingabe und -verwaltung aus dem Begriffskatalog dynamisch generieren. Durch

ein „proof of concept“ konnte gezeigt werden, dass diese innovative Entwicklung möglich ist. Die Umsetzung des GeoVal-DST bis zur Produktivversion erfolgt perspektivisch bis Ende 2021.

Für allgemeine Recherchen, Darstellungen, Visualisierungen, sowie einfache gezielte Datendownloads wird die Geotechnical Value Webapplication (GeoVal-Web) entwickelt. Der Nutzerkreis der hiermit angesprochen werden soll, setzt sich aus Fachanwendern und der fachlich interessierten Öffentlichkeit zusammen, die einen intuitiven Zugang zu den Daten suchen.

### 4.3 Datenmigration

Um alle Vorzüge von GeoValML auch für den bisherigen Altdatenbestand nutzen zu können, muss dieser aus bestehenden Datenbanken migriert werden. Aufgrund des deutlich anderen Schemas sind in der Regel umfangreiche Prozessierungen, Restrukturierungen und Zuweisungen notwendig. Die BAW hat sich dazu entschlossen die Migration der eigenen Labordatenbank mit Hilfe der Programmiersprache Python zu lösen. In Python werden vor allem die Packages Psycopg2, mit welchem SQL-Abfragen und Befehle an die Datenbank gerichtet werden können, und Pandas, zur internen Datenverwaltung, verwendet.

Das in Abbildung 2 skizzierte Migrationsschema stellt sich wie folgt dar:

Zunächst werden alle Daten der drei Bestandsdatenbanken geladen. Danach erfolgt die Datenprozessierung und -zuweisung. Vor dem Schreiben in die Zieldatenbank wird eine Fremdschlüsselprüfung durchgeführt. Ist diese erfolgreich, werden die Daten geschrieben. Falls nicht, werden die fehlenden Fremdschlüssel in eine EXCEL-Tabelle geschrieben. Dort können diese nach Tabellen sortiert analysiert werden. Nach der Fremdschlüsselprüfung und Korrektur kann das Migrationsprogramm neu gestartet werden.

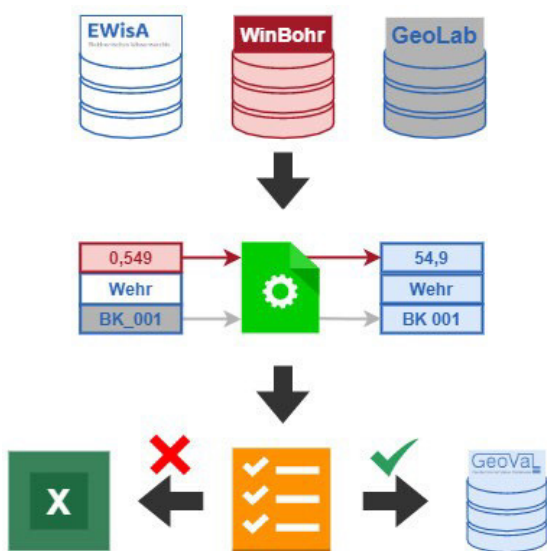


Abbildung 2: Migrationsschema

## Literaturverzeichnis

- [1] Bundesministerium für Wirtschaft und Energie. Geologiedatengesetz.  
<https://www.bmwi.de/Redaktion/DE/Artikel/Service/geologiedatengesetz.html>,  
03.05.2021.
  
- [2] Bundesministerium des Innern, für Bau und Heimat. E-Government-Gesetz.  
<https://www.bmi.bund.de/DE/themen/moderne-verwaltung/e-government/e-government-gesetz/e-government-gesetz-node.html>, 03.05.2021.
  
- [3] GO FAIR International Support and Coordination Office. FAIR Principles – GO FAIR.  
[www.go-fair.org/fair-principles/](http://www.go-fair.org/fair-principles/), 29.04.2021.

## Autoren

Dominik Stütz  
Bundesanstalt für Wasserbau  
Kußmaulstraße 17, 76187 Karlsruhe  
E-Mail : dominik.stuetz@baw.de

Eberhard Kunz  
Bundesanstalt für Wasserbau  
Kußmaulstraße 17, 76187 Karlsruhe  
E-Mail: eberhard.kunz@baw.de