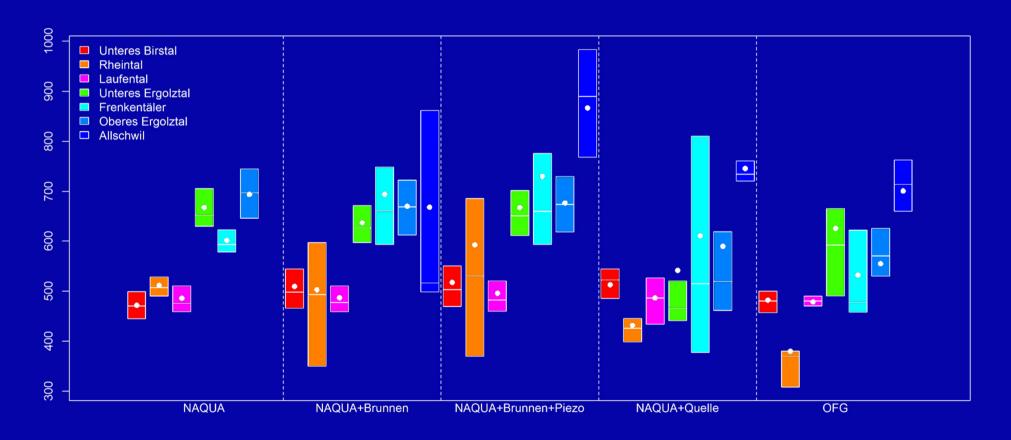
Datenhaltung und Auswertung von Grundwasserdaten



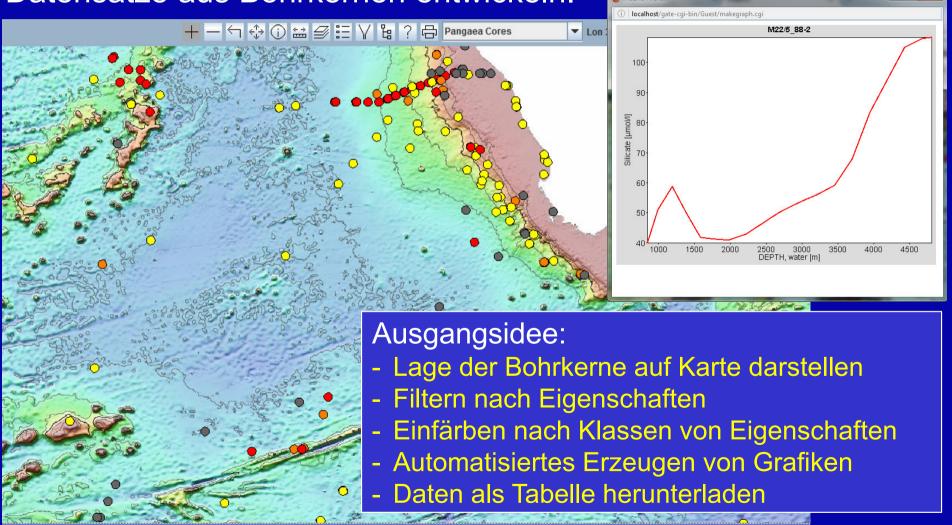
Christian Gmünder



Teil 1: Anforderungen an Datenhaltung

Als Industriepartner im Marie-Curie-Projekt «Gateways» sollte die Simultec eine Plattform zur Visualisierung paläoklimatischer

Datensätze aus Bohrkernen entwickeln.



Auswertung bestehender Daten

Die Datenplattform PANGAEA stellt Paläoklimatische Datensätze zum Download zur Verfügung



Eine Suche im Rechteck von -90° bis 0° Breite und -30° bis 60° Länge ergab folgende Resultate:

34'000 Datensätze

enthaltend:

- 90'000 geographische Objekte
- 48'000 davon befanden sich effektiv im Suchrechteck
- 20'000 unterschiedliche Parameterbezeichnungen
- 3'900 unterschiedliche Einheiten
- → Daten sind in Form einzelner Datenblätter abgelegt
- → Vorschriften zum Datenformat sind nicht sehr streng



Flache Datenstruktur (Tabellendaten)

Name	Lat	Lon	Depth	Parameter	Unit	Value	
DML03C98_09	-74.499	1.960	15.88	Phosphate	[µmol/l]	NV	
DML03C98_09	-74,499	1,960	842.3	Phosphate	[µmol/l]	2.34	
DNL03C98_09	-74.499	1.960	990.61	Phosphat	[µmol/l]	>2.35	
DML03C98_09	-74.499	1.960	1187.21	Phosphate	µmol/l	2.33	

Beispielsdatensatz

Tabellendaten weisen in 99% der Fälle folgende «Fehler» auf

- Schreibfehler
- Unterschiedliche Sprachen und Ländereinstellungen
- Unterschiedliche Schreibweisen für das Gleiche
- Schriftzeichen in Zahlenfeldern

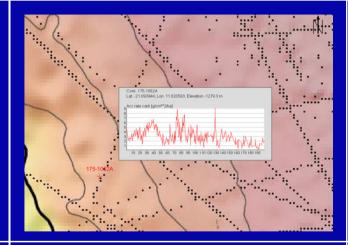
Anforderung an die Datenstruktur

Visualisierung

Visualisierung Anzeige der Daten in ihrem geographischen Kontext Filtern und Klassieren der Daten nach einer Eigenschaft

Automatisierte Visualisierung von Daten, Datenvergleich





Jeder Messwert muss einer bestimmten Koordinate zugeordnet werden können Die Anzahl der Parameter, Einheiten, Bohrmethoden, ... muss eine überschaubare sein Daten müssen in einer strukturierten Art abgelegt werden.

Anforderung



Relationale Datenbank

- Die gleiche Information wird nur einmal gespeichert
- Für die einzelnen Felder wird ein Datentyp festgelegt

BID	Name	Lat	Lon	
1	DML03C98_09	-74.499	1.960	
2	DML05C98_07	-74.499	1.960	
3	DML03C98_09	-74.499	1.960	

PID	Name	
1	Calcium	
2	Phosphate	
3	Nitrate	

r

	UID	Name	
n	26	mg/l	
	27	µmol/l	
	28	°C	

1 1/1

BID	PID	UID	Depth	Value	
1	2	27	15.88	2.27	
1	2	27	842.3	2.34	
1	2	27	990.61	2.35	
1	2	27	1187.21	2.33	

Resultierende Datenbank für die Bohrkerndaten 7

ProfileId

DATABASE

IsMethodOf

SampleMethodId

DataTypeId



Cluster

ClusterId Name Description

DataSet

DatasetId ClusterId Name Lat Lon Zco Length Project DataTypeId SampleMethodId Description KevWords SampleDate UploadDate PrimaryAuthorld Unpublished Ownerld ReadGroupId WriteGroupId

GateUser

Userld UserName Organisation **FirstName** LastName Email

SampleMethod

DataType

DataTypeId

Name

SampleMethodId Name Description

UserGroup

GroupId Name GroupType

Membership

Userld GroupId

Profile

ClusterId Name Description MethodId SubMethodId Instrument DataSetId PrimarySourceld UploadDate Unitld DistanceUnitId ErrorUnitId ParameterId PrimaryAuthorId Unpublished

AgeModel

AgeModeld DataSetId Name Description AgeUnitId DepthUnitId ErrorUnitId

Ownerld

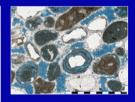
ReadGroupId

Pub

DataSetId Sourceld

Sample

ProfileId Distance SampleValue SampleError Note



IsSubOf

MethodId SubMethodId



Parameter

AgeValue

AgeModelId

Depth

AgeError

Sourceld

Citation

Abstract

Contact

Journal **Pages** Doi

PublicationYear

Title

Age

Method

SubMethod

SubMethodId

MethodId

Description

Name

Name

ParameterId Name ShortName DefaultUnitId

Unitd Name ShortName UnitTypeId

Unit

UnitType

UnitTypeId Name

Source **AutorShip**

Authorld Sourceld

Author

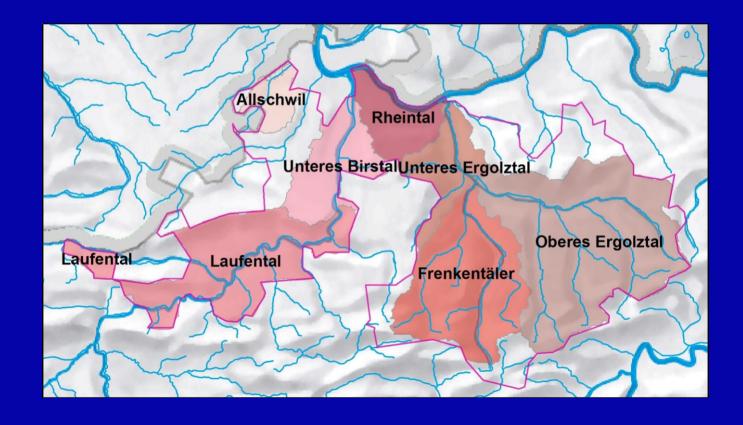
Authorld Name Contact



Teil 2: Auswertung von Grundwasserdaten

Grundwasserdaten des Kantons Basel Landschaft

- Der Kanton Basel Landschaft verfügt über eine grosse Menge an Daten zur Grundwasserqualität
- Die Daten stammen vom BAFU (Naqua), Trinkwasserversorgungen, Überwachungskampagnen und Altlastenuntersuchungen
- Geologie,
 Besiedlung,
 Industrie,
 Gewässer,
 Landwirtschaft
 sind lokal sehr
 unterschiedlich



Aufgabenstellung

Ausgangslage

- Es existiert bereits eine strukturierte Datenbank
- Der Kanton hat selbst einige Auswertungen in R programmiert

Unsere Aufgaben

- Die Auswertungen sollen ergänzt und erweitert werden
- Mehrere Auswertungen sollen automatisiert erzeugt und zu einem Bericht zusammengestellt werden
- Es soll eine grafische Benutzeroberfläche erstellt werden

Randbedingungen

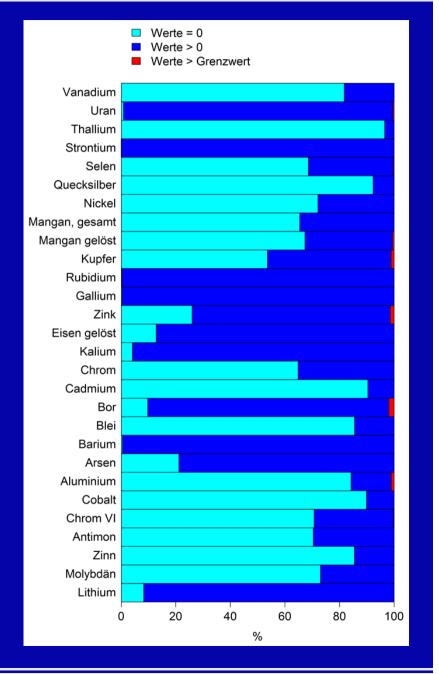
- Programmierung in R (Weiterverwendung durch Kanton)
- Nicht zu viel Aufwand für Benutzeroberfläche



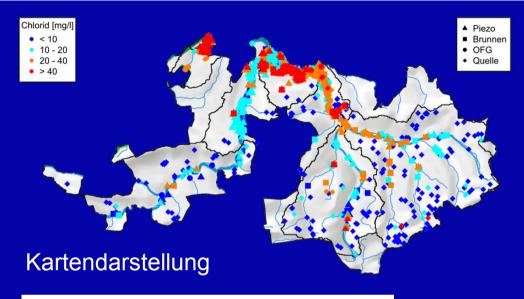
Übersicht gewinnen

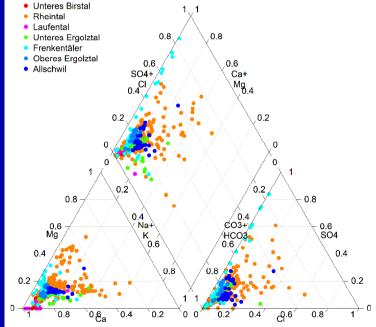
Die vorhandenen Daten wurden zuerst mit Übersichtsdarstellungen analysiert

- Viele Werte kleiner als Nachweisgrenze
- Werte können sich je nach Ursprung um Grössenordnungen unterscheiden (Trinkwasserüberwachung, Altlasten)
- Analyse führte zu einer Bereinigung der Datenbank (Fehler bei der Eingabe, falsche Einheiten, …)
- Aufgrund der grossen Streuung konnte der interessierende Bereich manchmal nicht dargestellt werden

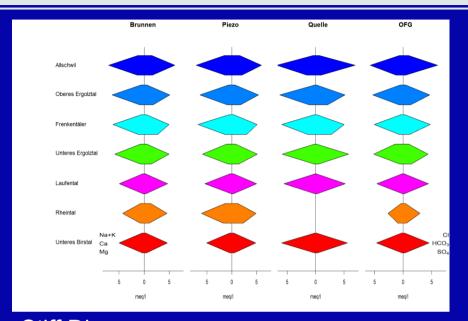


Darstellungsbeispiele

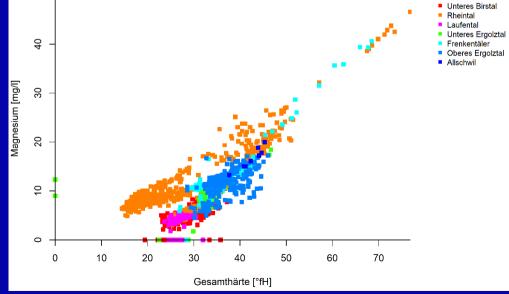




Piper Diagramm

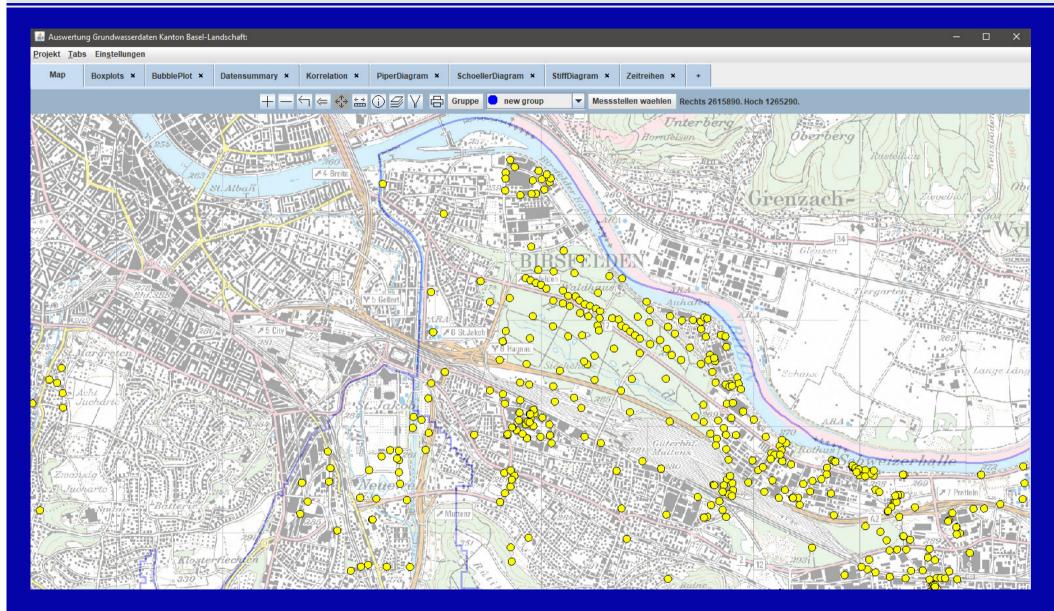






Korrelation

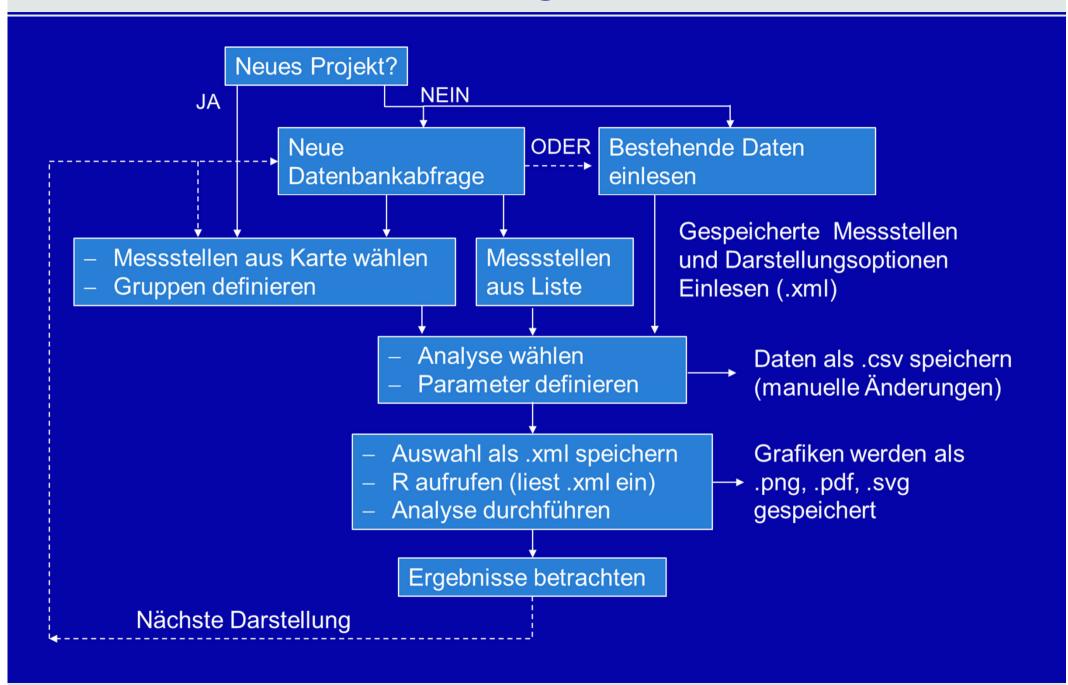
Grafische Benutzeroberfläche



Benutzeroberfläche in Java aus Projekt «Gateways» wurde erweitert



Workflow der Bedienung



Resultate

Es wurden zwei Berichte erstellt:

- «Geogene Einflüsse auf die Grundwasserqualität»
- «Antropogene Einflüsse auf die Grundwasserqualität»

Das automatische Erzeugen von Berichtsdarstellungen ist nicht in jedem Fall möglich

- Es muss eine Auswahl getroffen werden, da nur wenige Darstellungen interessant genug sind
- Teilweise sind zusätzliche, manuell erzeugte Darstellungen notwendig

Das Instrument eignet sich vor allem für folgende Aufgaben

- Qualitätskontrolle der eingegebenen Daten
- Übersicht über bestimmte Parameter, Gebiete gewinnen
- Die Bedienungsoberfläche erleichtert die Arbeit mit den Daten



Danke für die Aufmerksamkeit

Projektbeteiligte:

- Kanton Basel Landschaft, Amt für Umwelt und Energie
 - Dr. Adrian Auckenthaler
 - Dr. Dominik Bänninger
- Simultec AG
 - Dr. Ulla Heikkilä
 - Dr. Wei Li
 - Christian Gmünder

The work described in this presentation has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 – Marie-Curie ITN, under grant agreement n° 238512, GATEWAYS project.