

Modulbeschreibung

Geoprocessing

Allgemeine Informationen

Anzahl ECTS-Credits

3

Modulkürzel

TSM_GeoProc

Version

31. März 2016

Modulverantwortliche/r

Martin Christen, FHNW

Sprache

	Lausanne	Bern	Zürich
Unterricht	<input type="checkbox"/> E <input type="checkbox"/> F	<input type="checkbox"/> D <input type="checkbox"/> E <input type="checkbox"/> F	<input checked="" type="checkbox"/> D <input type="checkbox"/> E
Unterlagen	<input type="checkbox"/> E <input type="checkbox"/> F	<input type="checkbox"/> D <input type="checkbox"/> E <input type="checkbox"/> F	<input checked="" type="checkbox"/> D <input checked="" type="checkbox"/> E
Prüfung	<input type="checkbox"/> E <input type="checkbox"/> F	<input type="checkbox"/> D <input type="checkbox"/> E <input type="checkbox"/> F	<input checked="" type="checkbox"/> D <input type="checkbox"/> E

Modulkategorie

- Erweiterte theoretische Grundlagen
- Technisch-wissenschaftliche Vertiefung
- Kontextmodule

Lektionen

- 2 Vorlesungslektionen und 1 Übungslektion pro Woche
- 2 Vorlesungslektionen pro Woche

Kurzbeschreibung / Absicht und Inhalt des Moduls in einigen Sätzen erklären

Nachdem frühere Studien gezeigt hatten, dass 80% aller Entscheide im Geschäfts- und Privatleben auf Geodaten beruhen, gilt diese heute vermehrt auch für Big Data. Moderne Programmiersprachen wie Python haben sich dabei zu einem einfach zu erlernenden und zugleich mächtigen Werkzeug zur Prozessierung dieser Daten entwickelt. So ist Python die am häufigsten eingesetzte Programmiersprache für die Verarbeitung geografischer Daten, insbesondere wegen dessen Einfachheit, der umfangreichen Standardbibliothek und der Vielzahl an Erweiterungen.

Python wurde beispielsweise beim Space Shuttle Mission Design oder beim Prozessieren von Bildern des Hubble Space Telescope eingesetzt. Python wird aber auch bei hoch skalierbaren Web-Applikationen, wie beispielsweise Youtube oder bei der internen Google-Infrastruktur eingesetzt.

Parallel dazu hat sich die Verfügbarkeit von (Geo-)Daten im letzten Jahrzehnt drastisch verändert. So gibt es immer mehr Datenquellen – insbesondere auch dank Open Data Initiativen. Des Weiteren gibt es eine zunehmende Anzahl mobiler, internetfähiger Geräte welche es ermöglichen, von überall und zu jeder Zeit auf Daten zuzugreifen. Die Verfügbarkeit von Speicher- und Rechenkapazität durch Cloud-Computing Dienste hat dazu beigetragen, dass neue datenzentrische Dienstleistungen insbesondere auf mobilen Geräten und im Webbrowser möglich werden.

Eine grosse Herausforderung ist heute jedoch die Explosion der Datenmengen. Traditionelle Methoden zur Datenhaltung, Datenverarbeitung und Datenauswertung mittels relationaler Datenbanken und SQL waren jahrzehntelang der Weg um mit grossen Datenmengen zu arbeiten. Seit einigen Jahren gibt es jedoch Alternativen wie beispielsweise Hadoop, NoSQL oder Map-Reduce, welche innerhalb ihres Anwendungsgebietes viel effizienter damit umgehen können.

Dieses Modul zeigt die Einsatzmöglichkeiten von Python zur (Geo-)Datenprozessierung, zum Umgang mit Big Data, zum Datenhandling und Analyse und im Cloud-Computing. Dieses Modul zeigt die Einsatzmöglichkeiten von Python zur (Geo-)Datenprozessierung, zum Umgang mit Big Data, zum Datenhandling und Analyse und im Cloud-Computing, sowie die Kombinierbarkeit der einzelnen Bausteine zu einem gesamten Workflow.

Im ersten Teil wird die Programmiersprache Python eingeführt, dabei wird auch auf Python-Module für den Umgang mit Daten und dessen Verarbeitung eingegangen.

Im zweiten Teil wird auf die Thematik Big Data, Datenanalyse und Cloud Computing eingegangen.

Ziele, Inhalt und Methoden

Lernziele, zu erwerbende Kompetenzen
Fachliche Ziele:

- Erlernen der Programmiersprache Python und des Jupyter Notebooks
- Verwenden von Tools & Python-Modulen aus dem Bereich der Geodatenverarbeitung
- Umgang mit Python-Modulen im Bereich Big Data, NoSQL und Map-Reduce
- Die Fähigkeit erlangen, eigene Programme zu entwickeln, um Datensätze zu analysieren

- Prozessierung von Datensätzen in der Cloud

Methodische Ziele:

- Lösungskonzepte erarbeiten, vergleichen und präsentieren, erarbeitete Lösungen strukturieren und dokumentieren, Online-Foren finden und zur Problemlösung nutzen

Modulinhalt mit Gewichtung der Lehrinhalte

- Teil 1: Python & (Geo-)daten & Geo-Bibliotheken
 - Einführung in die Programmiersprache Python mit dem Jupyter Notebook
 - Einführung zu Numpy und Pandas
 - Einführung zu Raumbezug und Geographischen Daten (Referenz-Systeme, Projektionen, Transformationen etc.)
 - Prozessierung von räumlichen Daten (Lesen, Erstellen und Verändern von Raster- und Vektordaten)
 - Analyse und Visualisierung von räumlichen Daten
 - Arbeiten mit räumlichen Datenbanken (Import, Export, Datenmanipulation)
 - Verarbeitung von (räumlichen / OGC) Webdiensten
- Teil 2: Big Data Processing & Cloud Computing
 - Einführung Big Data
 - Big Data Datenmodell (Compute, Modellierung, Speicherung)
 - Datenspeicherung sehr grosser Datensätze
 - Hadoop (Python API, HDFS, Map-Reduce, Verarbeitung im Cluster)
 - NoSQL Datenbanksysteme (Anwendungsorientiert, Fokus Geodaten)
 - Skalierung mittels Cloud Computing und parallel Computing

Lehr- und Lernmethoden

- Vorlesungen (2 Lektionen pro Woche)
- Übungen (1 Lektion pro Woche)

Voraussetzungen, Vorkenntnisse, Eingangskompetenzen

Grundlegende IT-Kenntnisse

Interesse am Umgang mit Geodaten

Interesse an Cloud-Umgebungen und Prozessierung sehr grosser Datensätze

Bibliografie

- Wes McKinney, 2012. Python for Data Analysis: Data Wrangling with Pandas, Numpy and IPython.
- Erik Westra, 2013. Python Geospatial Development. Second Edition. Packt Publishing, Birmingham
- Joel Lawhead, 2013. Learning Geospatial Analysis with Python. Packt Publishing
- Chris Garrard, 2015. Geoprocessing with Python, Manning, ISBN: 9781617292149.

Leistungsbewertung**Zulassungsbedingungen für die Modulschlussprüfung (Testatbedingungen)**

Aktive Teilnahme am Unterricht (min. 80% Anwesenheit) und lösen der Übungen.

Schriftliche Modulschlussprüfung

Prüfungsdauer : 120 Minuten

Erlaubte Hilfsmittel: 5 Seiten (Grösse A4) handgeschriebene Zusammenfassung