Module Description, available in: EN

Geoprocessing

General Information

Number of ECTS Credits

3

Module code

TSM_GeoProc

Valid for academic year

2020-21

Last modification

2019-09-22

Coordinator of the module

Martin Christen (FHNW, martin.christen@fhnw.ch)

Explanations regarding the language definitions for each location:

- Instruction is given in the language defined below for each location/each time the module is held.
- Documentation is available in the languages defined below. Where documents are in several languages, the percentage distribution is shown (100% = all the documentation).
- The examination is available 100% in the languages shown for each location/each time it is held.

	Lausanne			Lugano	Zurich		
Instruction					**X** E 100%		
Documentation					**X** E 100%	**X** D 70-80%	
Examination					**X** E 100%		

Module Category

TSM Technical scientific module

Profile Recommendations

This module is recommended for the following profiles:

	Zurich
Profile	Geo

Lessons

2 lecture periods and 1 tutorial period per week

Entry level competences

Prerequisites, previous knowledge

Basic IT knowledge
Interest in dealing with geodata
Interest in cloud environments and processing of very large data sets

Brief course description of module objectives and content

After studies had shown that 80% of all decisions in business and private life are based on geodata, this now also applies increasingly to Big Data. Modern programming languages such as Python have developed into an easy-to-learn and powerful tool for processing this data. Python is the most frequently used programming language for processing geographical data, especially because of its simplicity, the extensive standard library and the large number of extensions.

Python was used, for example, for Space Shuttle Mission Design or for processing images from the Hubble Space Telescope. Python is also used for highly scalable web applications such as YouTube or the internal Google infrastructure.

At the same time, the availability of (geospatial) data has changed drastically in the last decade. There are more and more data sources - especially thanks to Open Data initiatives. In addition, there is an increasing number of mobile, Internet-enabled devices that make it possible to access data from anywhere and at any time. The availability of storage and computing capacity through cloud computing services has helped to make new data-centric services possible, especially on mobile devices and in web browsers.

A major challenge today, however, is the explosion of data volumes. Traditional methods for data storage, data processing and data evaluation using relational databases and SQL have been the way to work with large amounts of data for decades. For some years, however, there have been alternatives such as Hadoop, NoSQL or Map-Reduce, which can handle this much more efficiently within their area of application.
This module shows the possible uses of Python for (geo)data processing, for handling big data, for data handling and analysis and in cloud computing.
This module shows the application possibilities of Python for (geo) data processing, for handling big data, for data handling and analysis and in cloud computing, as well as the combinability of the individual building blocks to an entire workflow.

In the first part, the programming language Python is introduced, and Python modules for handling and processing data are discussed.

The second part deals with big data, data analysis, and cloud computing.

Aims, content, methods

Learning objectives and competencies to be acquired

Technical goals:

- Learning the programming language Python and the Jupyter Notebook
- Using Tools & Python modules from the field of geo data processing
- Dealing with Python modules in the area of Big Data, NoSQL and Map-Reduce
- Gain the ability to develop your own programs to analyze data sets
- Processing of data records in the cloud

Methodological objectives:

- Develop, compare and present solution concepts, structure and document developed solutions, find online forums and use them to solve problems.

Module content with weighting of different components

Part 1: Python & (geo-)data & geo-libraries

- Introduction to the programming language Python with the Jupyter Notebook
- Introduction to Numpy and Pandas
- Introduction to spatial reference and geographical data (reference systems, projections, transformations etc.)
- Processing of spatial data (reading, creating and modifying raster and vector data)
- Analysis and visualization of spatial data
- Working with spatial databases (import, export, data manipulation)
- Processing of (spatial / OGC) web services

Part 2: Big Data Processing & Cloud Computing

- Introduction Big Data
- Big Data data model (compute, modeling, storage)
- Data storage of very large data sets
- Hadoop (Python API, HDFS, Map-Reduce, Cluster processing)
- NoSQL database systems (application-oriented, focus on geodata)
- Scaling using cloud computing and parallel computing

Teaching and learning methods

Lectures (2 lessons per week)
Exercises (1 lesson per week)

Literature

- 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.

Assessment

Certification requirements

Module uses certification requirements

Certification requirements for final examinations (conditions for attestation)

Active participation in lessons (min. 80% attendance) and solving exercises.

Basic principle for exams

As a rule, all standard final exams are conducted in written form. For resit exams, lecturers will communicate the exam format (written/oral) together with the exam schedule.

Standard final exam for a module and written resit exam

Kind of exam

Written exam

Duration of exam

120 minutes

Permissible aids

Aids permitted as specified below:

Permissible electronic aids

No electronic aids permitted

Other permissible aids

Teaching materials (Script/Jupyternotebook) printed out

Special case: Resit exam as oral exam

Kind of exam

Oral exam

Duration of exam

30 minutes

Permissible aids

Aids permitted as specified below:

Permissible electronic aids

Teaching materials (Script/Jupyternotebook) printed out

Other permissible aids

No other aids permitted