humanities.txt

Göttingen Summer School in Digital Humanities (August 2013)

Mike Kestemont (University of Antwerp) & Lars Wieneke (CVCE Luxembourg)

In the Humanities, scholars study the **products of the human mind**. These products are diverse (language, philosophy, visual arts, literature, law, etc.) but, typically, they tend to have an important **textual component**. It has become a platitude that with the ever increasing corpus of electronic data available, scholars can nowadays engage with texts on an unprecedented scale. **Innovative research** questions emerge – often of a diachronic nature – and scholars are trying to provide, new **data-driven answers** to age-old questions in the Humanities (cf. Culturomics). In the **Digital Humanities**, a lot of attention is currently invested in the development of computational methodologies to assist in this process.

In this Summer School, we will focus on **computational text analysis**: how can we make use of computation to enhance, and even alter our traditional approach of textual resources in the Humanities? Together, we will explore how **a little bit of scripting can make a huge difference** in our daily research practice. Throughout the summer school, we will make extensive use of the **Python** scripting language, which has recently become the de facto standard in scientific programming, especially computational linguistics. We will first look into basic concepts such as variables, function, loops, ... and explore why we need these to be able to automatically process digital texts. After an introduction to the **fundamentals** of Python, we will deepen our understanding of language with respect to **practical applications**, learning how to process XML documents and steer influential software packages like the *Natural Language Toolkit*, *Pattern* and other tools in the domain of Machine Learning and Topic Modeling. We will only work with real-world, open-source data, taken from resources like the *Google N-Grams Corpus* and *Gutenberg.org*.

This Summer School will be extremely hands-on, with half of the sessions devoted to making exercices, both in group and individually. Participants will additionally develop skills in online trouble-shooting. The program is aimed at all students and scholars in the Humanities, especially those with **no previous programming experience whatsoever**. Participants should bring a – not too ancient – laptop, running one of the major operating systems (Windows, Mac OS X or Linux). Installation instructions for a number of software packages will be sent out to the participants in the weeks before the workshop.