



BA/SA/Group/Lab:

Style Change Detection

Imagine that you are a teacher searching for evidence that will allow you to identify a cheating student. You have a feeling that the style of his writing has changed significantly within the text, but how can we quantify this?

Indeed, the following problem is relevant: Given an *unknown* text document, is it possible to decide whether the author style is consistent throughout the entire document or not? How (and how well) can we quantify author style and author style changes?

There has been a lot of research on this topic area (see authorship attribution). We may start by looking at some of the known strategies and their attribution capabilities; however, original approaches are very welcome. To evaluate our implementation, we may ultimately participate in the PAN shared tasks (<https://pan.webis.de/tasks.html>) and compete against international scientists from this specific domain.

Requirements: Programming experience (especially in Python) might be an advantage. The topic is broad and can be tailored to your previous knowledge and interests. There will be weekly meetings with your supervisor to discuss progress and open questions. For a group project, there might be the option to do an intense hackathon weekend to start off.

Interested? Please contact us for more details!

Contacts

- Roland Schmid: roschmi@ethz.ch, ETZ G94

