

Universität Koblenz-Landau

**Name:** Christoph Carl Kling

**Titel der Dissertation:** “Probabilistic Models for Context in Social Media: Novel Approaches and Inference Schemes”

(Deutscher Titel: “Neue Wahrscheinlichkeitsmodelle und Inferenz-Techniken für Kontextinformationen im World Wide Web”)

**Zusammenfassung:**

Thema meiner Dissertation ist die Erkennung von Mustern in Web-Dokumenten mit Metadaten. Schwerpunkt der Arbeit sind sogenannte *Topic Models*. Mithilfe von *Topic Models* können automatisch Themen in großen Dokumentensammlungen erkannt werden. Dazu wird das gemeinsame Auftreten von Wörtern in Dokumenten genutzt: Mengen an häufig gemeinsam auftretenden Wörtern werden als Themen (*Topics*) interpretiert.

Dokumente aus dem Web sind oft mit Metadaten wie etwa Zeitstempeln, Ortsinformationen oder Informationen über den Autor / die Autorin versehen. Metadaten beschreiben in vielen Fällen den Kontext, in dem Dokumente erstellt wurden. Diese Kontextinformationen können die Themen-Erkennung durch *Topic Models* verbessern oder sogar erst ermöglichen. Außerdem erlauben sie die Analyse von Zusammenhängen zwischen Kontext und Themen, etwa an welchen Orten welche Themen populär sind.

In meiner Arbeit werden neuartige *Topic Models* vorgestellt, die die Einbeziehung von (fast) beliebigen Kontextinformationen erlauben, messbar die Qualität der erkannten Themen verbessern und aufgrund ihrer Struktur eine effiziente Inferenz ermöglichen. Zudem sind die Parameter der Modelle interpretierbar und können auch komplexe Zusammenhänge zwischen Kontext und Themen aufdecken, die mit bisherigen Modellen nicht erkannt werden können.