



TECHNISCHE
UNIVERSITÄT
DARMSTADT

eCodicology



Algorithmen zum automatischen Tagging mittelalterlicher Handschriften

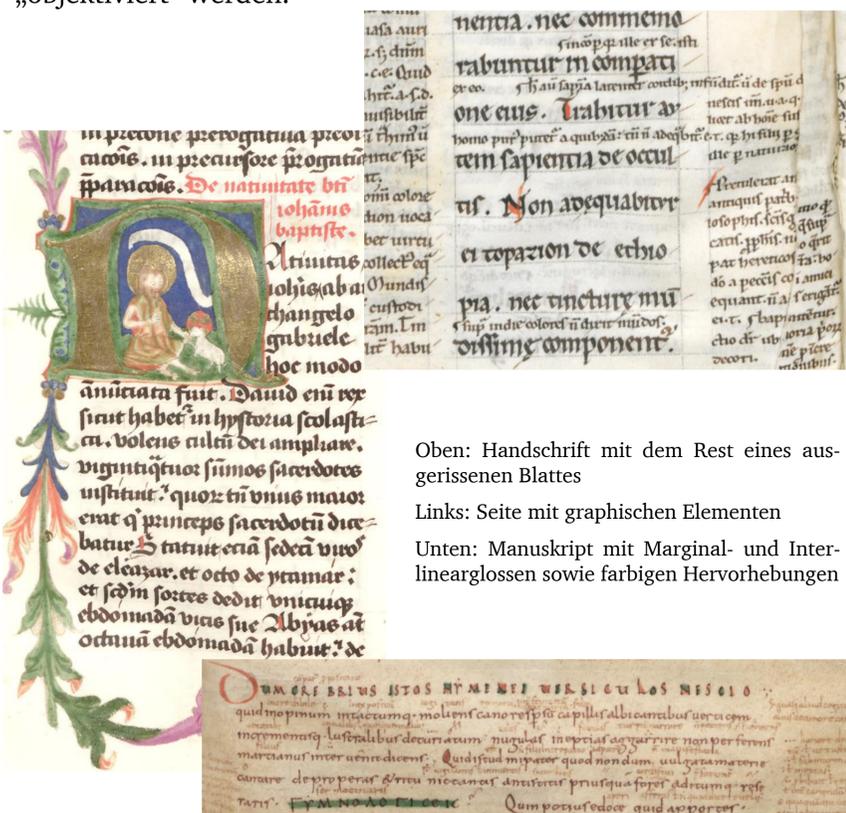
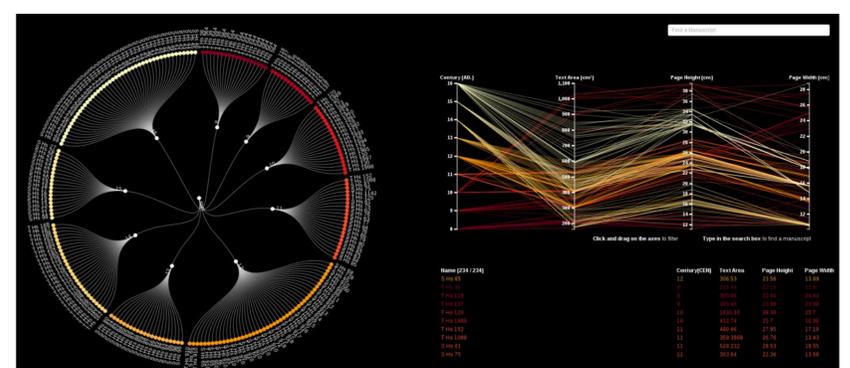


eCodicology nutzt den Fundus des digitalisierten mittelalterlichen Bestandes von rund 500 Handschriften aus der Benediktinerabtei St. Matthias in Trier, der im Projekt „Virtuelles Skriptorium“¹ mit bibliographischen Metadaten erschlossen wurde. Ziel von eCodicology ist die Entwicklung, Erprobung und Optimierung neuer informationstechnologischer Algorithmen, die makro- und mikrostrukturelle Layoutmerkmale der Handschriftenseiten automatisch erkennen und in den Metadaten als XML gemäß den TEI-Richtlinien ableiten. Auf dieser Datenbasis aufbauend können für viele geisteswissenschaftliche Disziplinen relevante Fragen nach Schreiber corpora, Schreibschulen, Handschriftenbezügen, Provenienzen und Ähnlichem auf empirischer Basis analysiert werden, der „subjektive“ Blick des Kodikologen kann dadurch „objektiviert“ werden.

```
<tei:physDesc><tei:objectDesc form="codex">
  <tei:supportDesc>
    <tei:material>Pergament</tei:material>
    <tei:watermark></tei:watermark>
  </tei:supportDesc>
  <tei:extent>
    <tei:measure type="leavesCount">44&#160;Bl.</tei:measure>
    <tei:measure type="pageDimensions">8</tei:measure>
  </tei:extent>
  <tei:condition>Das Manuskript ist infolge von Kriegseinwirkungen stark beschädigt. Die Schrift ist stark ausgewaschen und es gibt Textverluste.</tei:condition>
  <tei:layoutDesc>
    <tei:layout><tei:measure type="written"></tei:measure></tei:layout>
  </tei:layoutDesc>
  <tei:handDesc>
    <tei:handNote><tei:p>Frater Arnold Gernersdunck; Arnold de moys (aus Moers)</tei:p></tei:handNote>
  </tei:handDesc>
  <tei:decoDesc>
    <tei:decoNote><tei:p>Blatt 5r: Reiche Initiale auf Goldgrund in Pflanzenornament mit zwei Miniaturköpfen.</tei:p></tei:decoNote>
  </tei:decoDesc>
  <tei:bindingDesc>
    <tei:binding><tei:p>Prachteinband aus dem 17. Jh., Holzdeckel in braunem Leder mit reicher Pressung: Darstellungen aus dem Leben Christi, Brustbilder von Heiligen und Pflanzenornamente. Rücken später mit Schweinsleder erneuert.</tei:p></tei:binding>
  </tei:bindingDesc>
  <tei:accMat></tei:accMat>
</tei:physDesc>
```

Äußere Merkmale der Handschriften, die durch Bildverarbeitung und Mustererkennungsprozesse bestimmt werden können und deren statistische und qualitative Auswertung helfen kann, Handschriftengruppen nach Provenienz, Schreiberhänden oder Skriptorien zu identifizieren, sind unter anderem Seitengröße, Schriftraum, Spaltenzahl, verschiedene graphische Elemente und die Verhältnisse von Text- zu Bildraum. Die Messergebnisse werden in XML-Dateien mit den entsprechenden Bilddaten im TextGrid Repository² langzeitarchiviert.

Der in einem kooperierenden Projekt entwickelte Webdienst CodiLab³ erlaubt, die identifizierten Bildbereiche mit einer kodiko-logischen Ontologie zu verknüpfen und Fehler zu korrigieren; so lassen sich die Resultate terminologisch exakter bestimmen. Das in eCodicology entwickelte Werkzeug CodiVis gibt in einem radialen Baumdiagramm einen schnellen Überblick über alle vermessenen Handschriften und stellt in einem parallelen Koordinatensystem die Verteilung der verschiedenen Layoutmerkmale dar. CodiVis wird ab 2016 von DARIAH-DE⁴ zur Verfügung gestellt werden.



Oben: Handschrift mit dem Rest eines ausgerissenen Blattes

Links: Seite mit graphischen Elementen

Unten: Manuskript mit Marginal- und Interlinearglossen sowie farbigen Hervorhebungen

Weitere Informationen:

¹ Virtuelles Skriptorium St. Matthias: stmatthias.uni-trier.de

² TextGrid: textgrid.de ⁴ DARIAH-DE: de.dariah.eu

³ CodiLab: jochengraf.github.io/CodiLab/index.html



Kontakt:
Prof. Dr. Andrea Rapp
rapp@linglit.tu-darmstadt.de
www.eCodicology.org



GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung