

Algorithms for the Automatic Tagging of Medieval Manuscripts

# QuantiCod Revisited

## Neue Ansätze zur quantitativen Analyse mittelalterlicher Handschriftenbestände

Hannah Busch, Swati Chandna, Celia Krause, Philipp Vanscheidt

DHd 2015

Karl-Franzens-Universität Graz

26. Februar 2015



Bundesministerium  
für Bildung  
und Forschung



Kompetenzzentrum  
für elektronische Erschließungs-  
und Publikationsverfahren in  
den Geisteswissenschaften

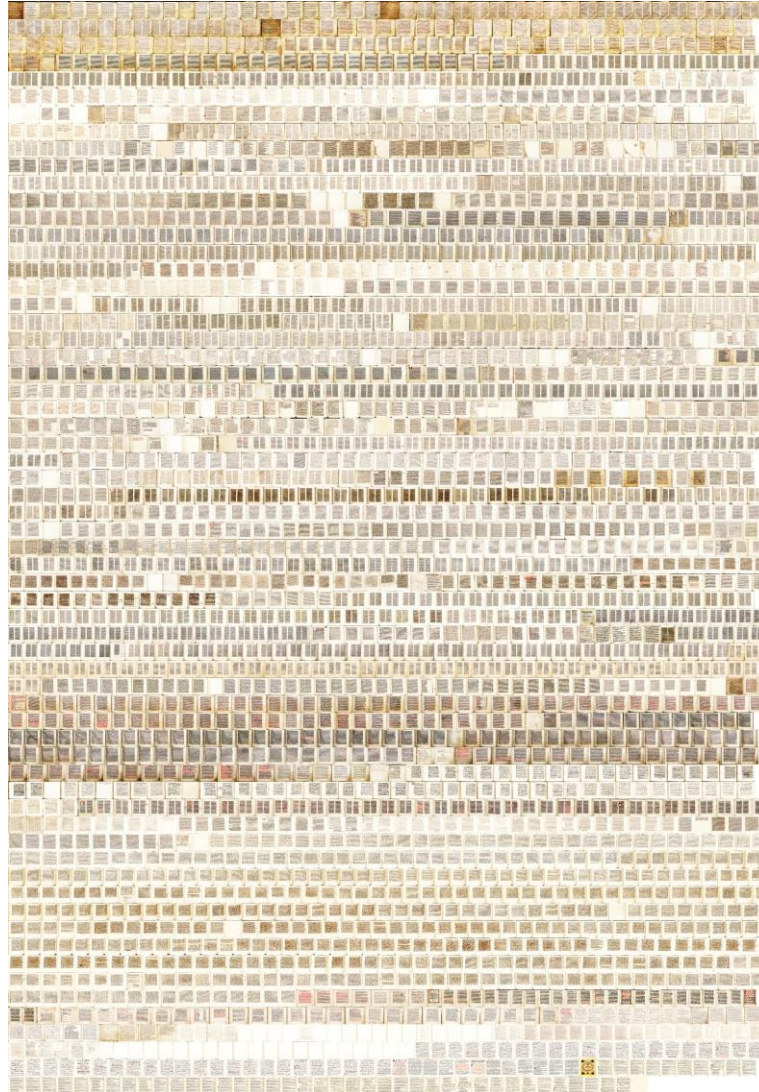
**KIT**  
Karlsruher Institut für Technologie



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

- Quantitative Kodikologie
- Untersuchungsgegenstände
- Komparativer Ausblick
- Anwendungsbeispiel „eCodicology“

# Quantitative Kodikologie



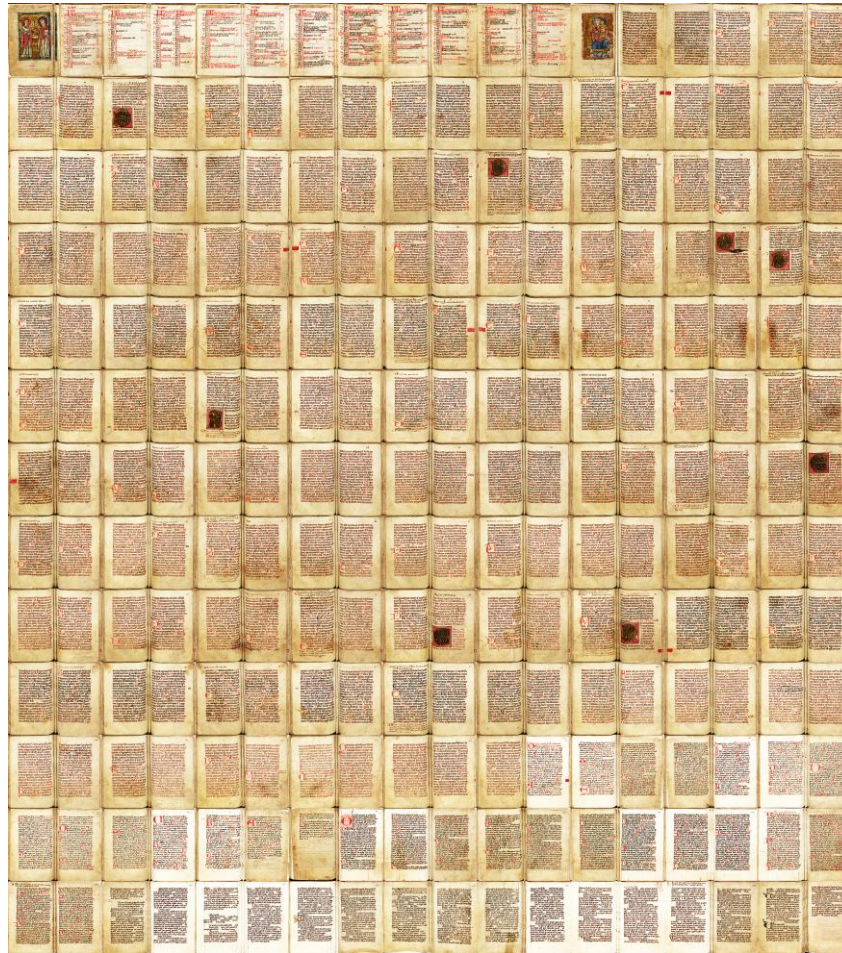
# Geschichte der quantitativen Kodikologie

- Seit etwa 1950: Die Kodikologie ist fester Bestandteil der historischen Hilfswissenschaften
- Etwa 1980–1995: Gruppe quantitativer Kodikologen am CNRS Paris.
- Erste Einsätze des Computers für die Datenverarbeitung.
- Seit etwa 2000: Weitreichende Digitalisierungsprogramme, verstärkter Computereinsatz.

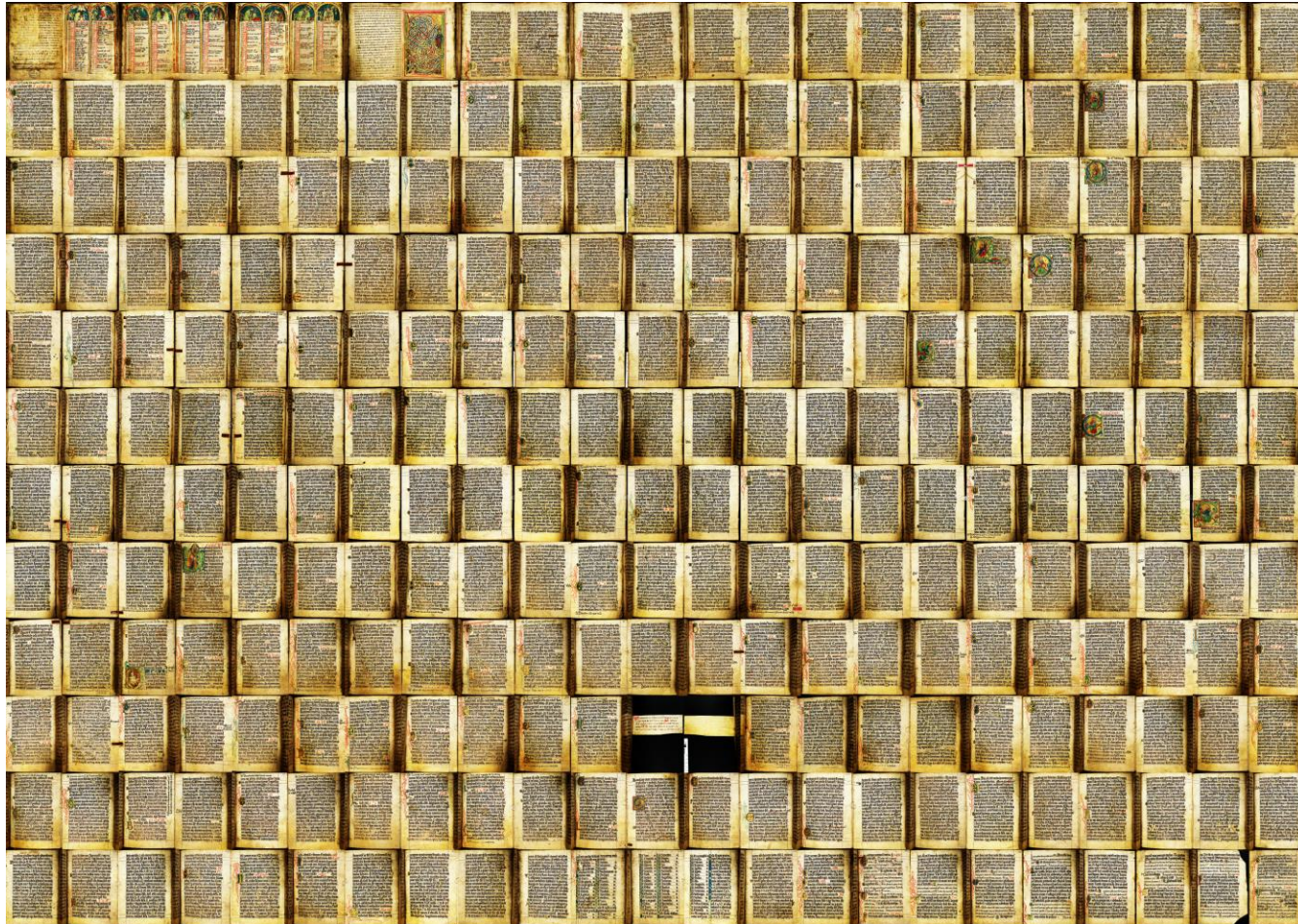
# Programm der quantitativen Kodikologie

- „Die Beobachtung des Banalen“ – Erkenntnis weit verbreiteter Effekte
- Verwendung statistischer Modelle (Korpusbildung, Variablennormierung, Signifikanzüberprüfungen)
- Verknüpfung modaler und kausaler Analysen (kommunikative Funktionen, technische Dispositive, sozioökonomischer Rahmen)

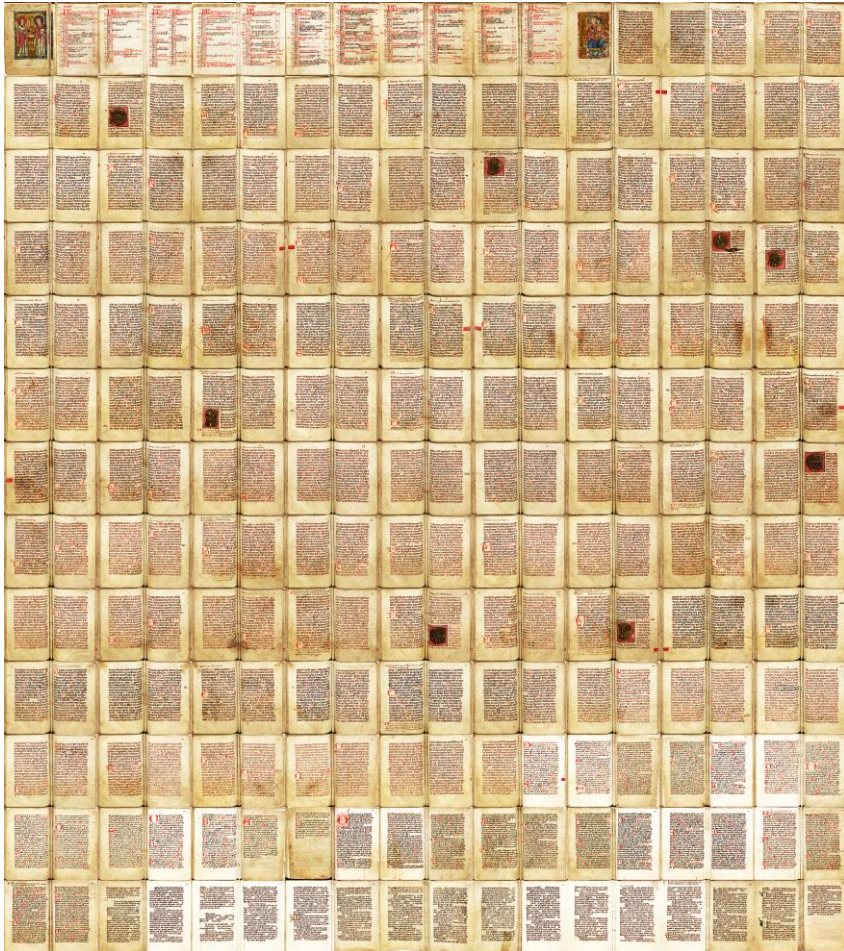
# Komparative Kodikologie



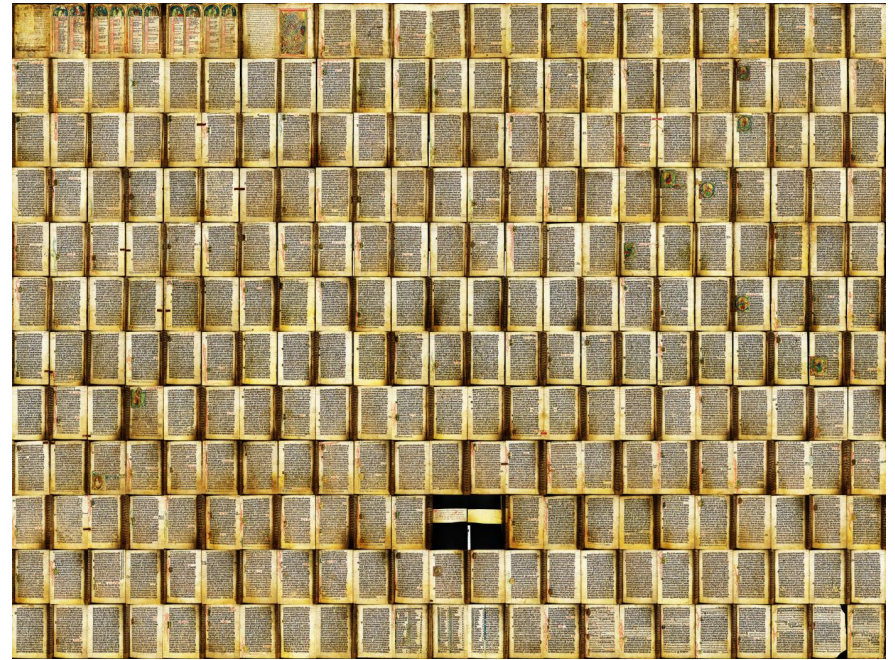
Trier



Wolfenbüttelese



Stadtbibliothek Trier, Hs. 429



HAB Cod. Guelf. 989

# Motivation



## Was ist eCodicology?

Ein vom BMBF gefördertes Projekt der

- Technischen Universität Darmstadt
- Trier Center for Digital Humanities
- Karlsruhe Institute for Technology

## Was will eCodicology?

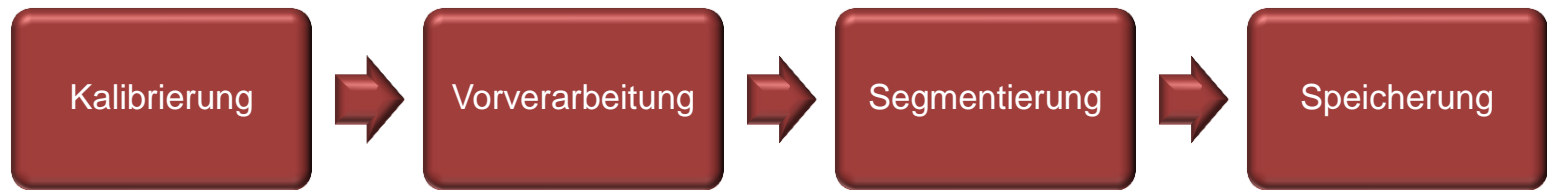
- Automatische Erkennung von makro-und mikrostrukturellen Gestaltungsmerkmalen
- Statistische Analyse reproduzierbarer Merkmale
- Erkenntnis von Oberflächlichkeiten und versteckten Beziehungen

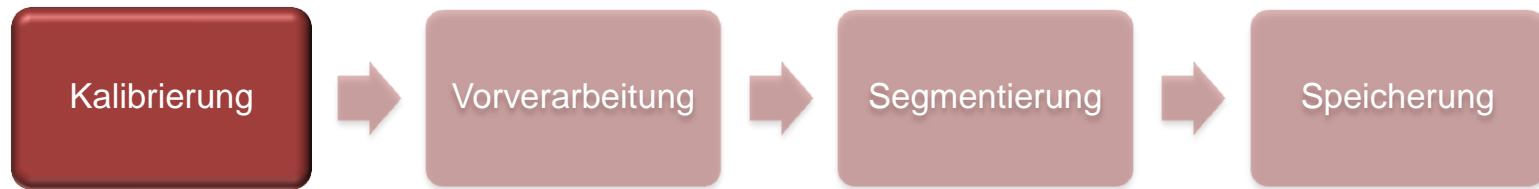


# Merkmalsextraktion

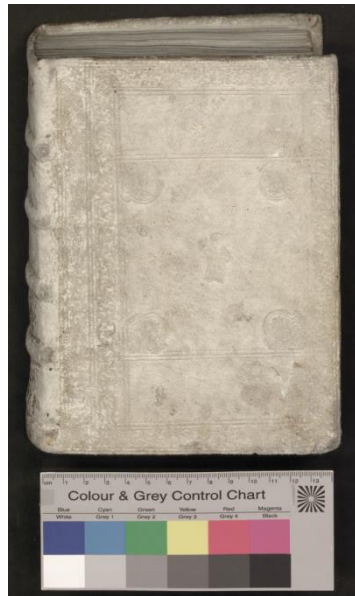
**Merkmal:** Meßbare Eigenschaft eines beobachteten Objekts







# Farb- und Größenkalibrierung



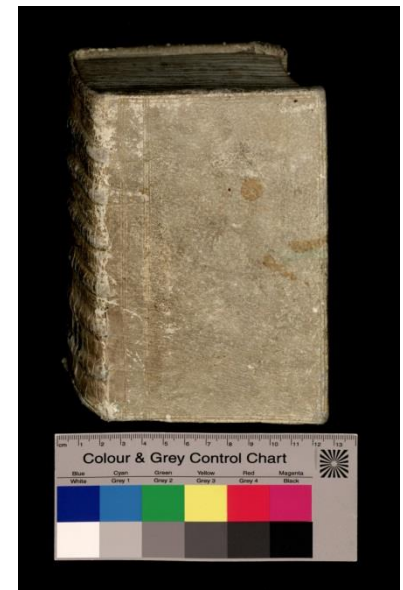
Scanner 1



Scanner 2



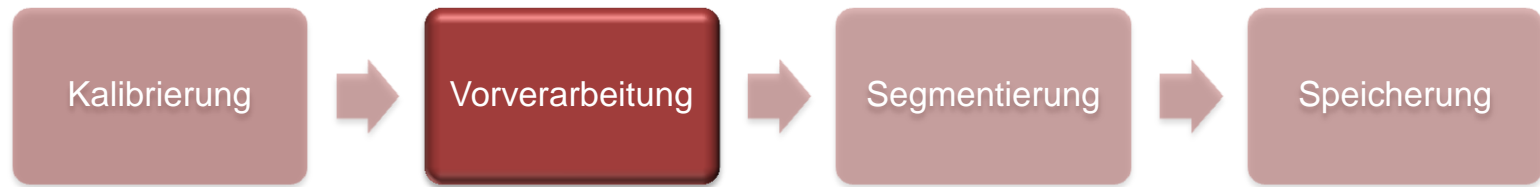
Scanner 1



Scanner 2

Vor der Kalibrierung

Nach der Kalibrierung

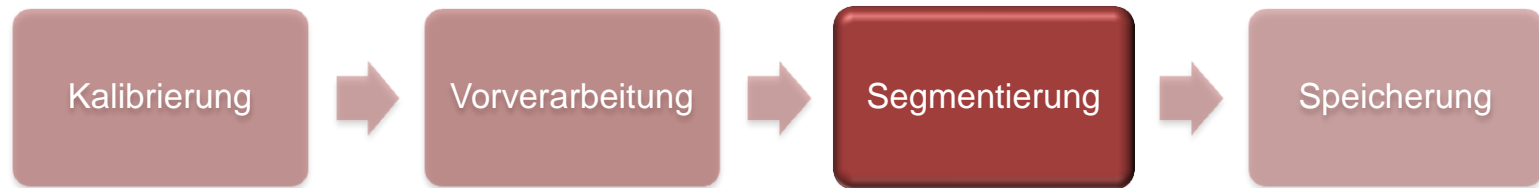


# Filter



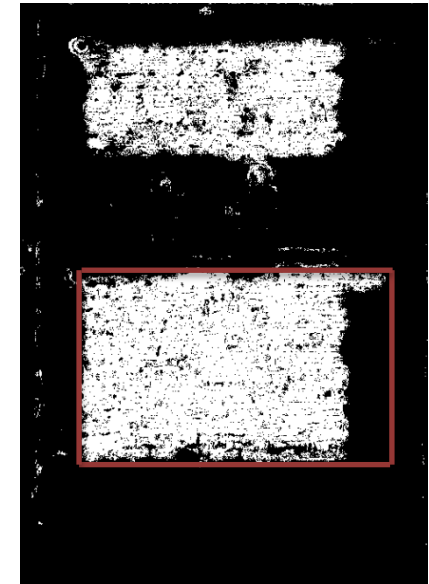
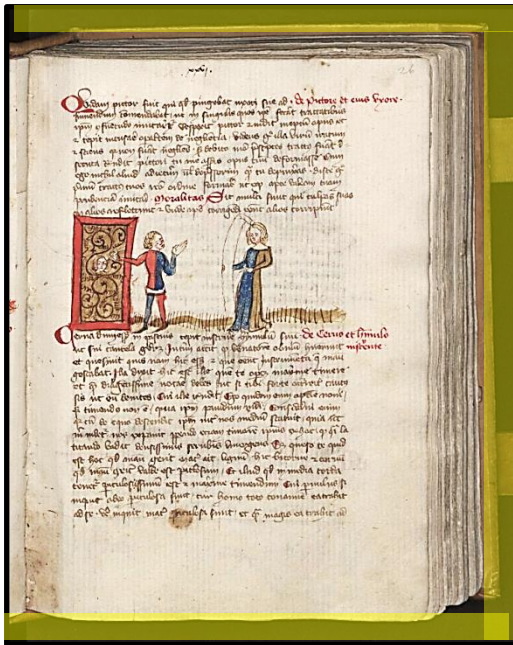
# Skalierung

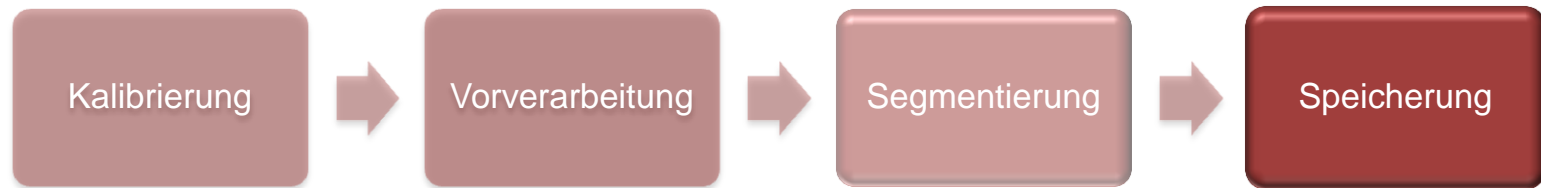




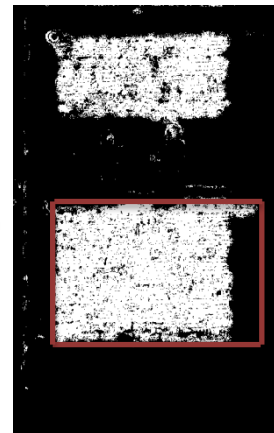
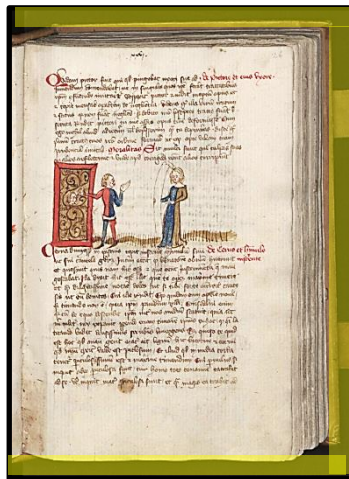
# Segmentierung

- Unterteilung des Bildes nach ähnlichen Eigenschaften über Trainings
- Bestimmung von interessanten Bereichen





# Speicherung



```

<xml>
...
<width>250
</width>
<unit>mm
</unit>
<height>550
</height>
<unit>mm
</unit>
...</xml>
  
```

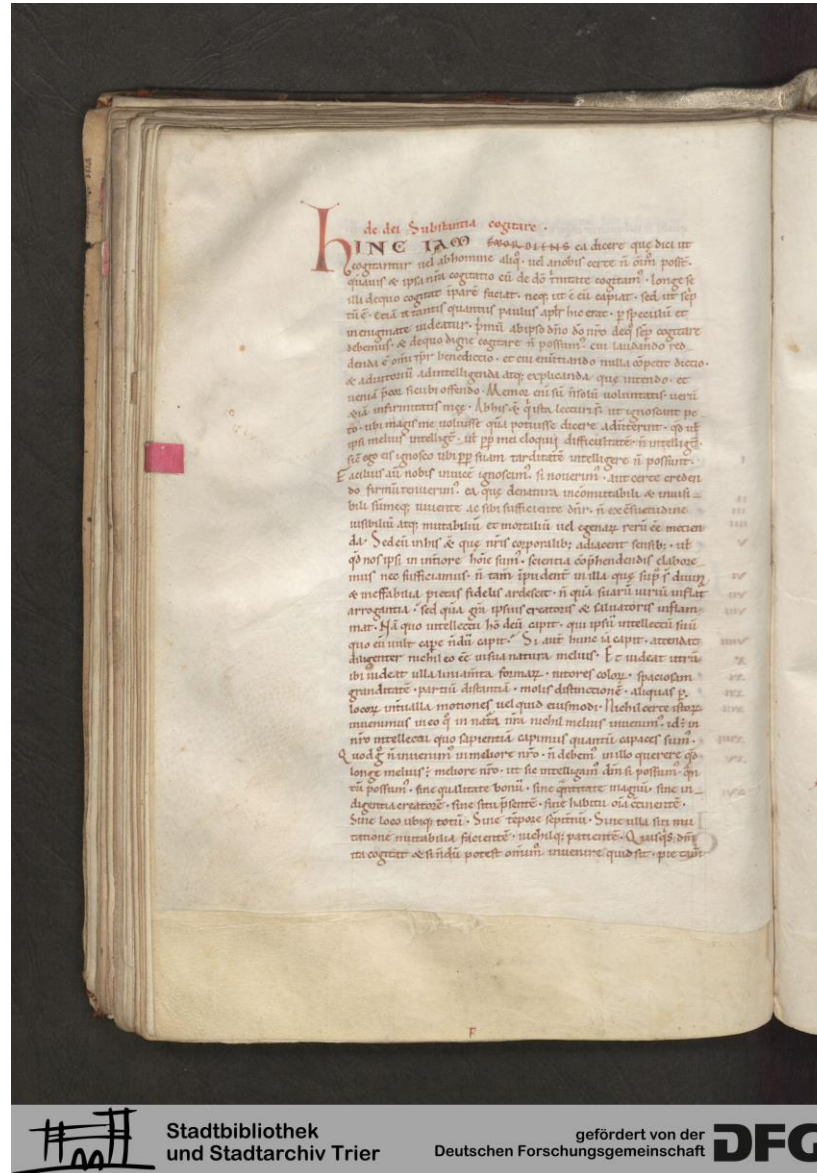
# Verlässlichkeit der Faksimiles

- Die Reproduktion ist nicht das Original
  - Größenverhältnisse
  - Farbechtheit
  - Dimensionalität
  - Artefakte, Rauschen, Fehler
- Gegenmittel:
  - Kontrolle bestimmter Variablen
  - Berechnung systematischer Abweichungen

## Ergebnisse: Einzelfall

- Durchschnittliche prozentuale Abweichung der Messungen bezogen auf die Seitenhöhe: 2,07.
- Durchschnittliche prozentuale Abweichung der Messungen bezogen auf die Seitenbreite: 9,43.

# Ergebnisse: Einzelfall



## Ergebnisse: Serie

- Durchschnittliche prozentuale Abweichung der Messungen bezogen auf die Seitenhöhe: 2,4%.
- Durchschnittliche prozentuale Abweichung der Messungen bezogen auf die Seitenbreite: 14,11%.
- Mittlere Abweichung der manuellen Höhenmessungen: 0,83 mm.
- Mittlere Abweichung der automatischen Höhenmessungen: 1,11 mm.
- Mittlere Abweichung der manuellen Breitenmessungen: 2,05 mm.
- Mittlere Abweichung der automatischen Breitenmessungen: 5,72 mm.

# Vorschlag für Richtlinien

1. Die durchschnittliche Abweichung zwischen manueller und automatischer Messung sollte nicht mehr als 2,5% betragen.
2. Jede Abweichung von über 5% kann als auffällig gelten.

Hannah Busch  
Universität Trier  
buschh@uni-trier.de

Philipp Vanscheidt  
Universität Trier/Technische Universität Darmstadt  
pvanscheidt@uni-trier.de

eCodicology kooperiert mit

