

Annotationen zur mittelalterlichen Buchgestalt

Celia Krause, Philipp Vanscheidt

TU Darmstadt



eCodicology in Kürze

Laufzeit: 2013-2016

Datenbasis: Virtuelles Skriptorium St. Matthias
(DFG-Projekt, 2010-2014)

ca. 500 digitalisierte mittelalterliche Handschriften
= etwa 170.000 Einzelbilder

Ziele:

- Detektion und Extraktion von formalen Layoutmerkmalen auf den gescannten Handschriftenseiten mittels Bilderkennungsalgorithmen (Ausdehnung, Anzahl und Position auf der Seite)
- Überführung der Messergebnisse in TEI-konformes XML
- statistische Auswertung und Visualisierung der Ergebnisse

Kooperation mit SemToNotes

CodiLab

CodiAnn: Bildannotationstool zur manuellen Korrektur und zur semantischen Anreicherung von automatischen Bildannotationen

CodiKOS: browser-basierter SKOS Editor für die Einbettung einer kodikologischen Ontologie (rdf-basiert)

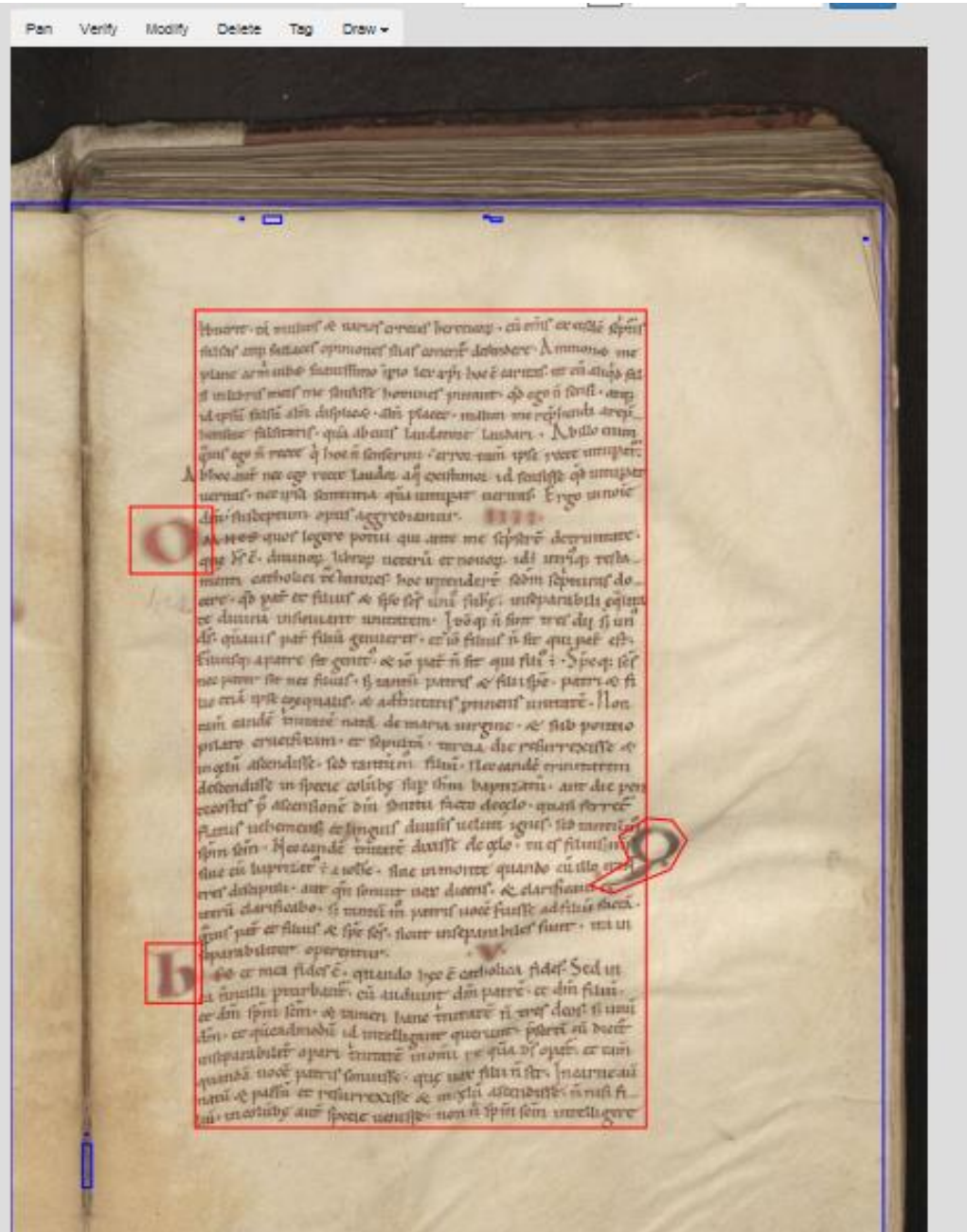
CodiQ: topologisches Retrievalsystem zur Analyse/Kontrolle von automatischen Bildannotationen

Tags Overview

- Illumination
- Body of page
- Initial
- Automatic annotations

Selected Annotation

CodiAnn



CodiKOS

Browse SKOS Concepts CodiKOS - CodiKOS (CodiKOS)

[View CodiKOS on GitHub »](#)

CodiKOS

0001 Physical description - Physikalische Beschreibung

-

0002 Dimensions - Maße

E.g. height and width of a leaf, dimensions of the written space the text block (by measuring the frame ruling); dimensions usually given in millimetres. - z.B. Höhe und Breite eines Blattes, Ausmaße des beschriebenen Raumes auf dem Blatt (durch Messen der Rahmenlinierung); gewöhnlich in Millimeter angegeben.

0003 Pricking - Einstechen

The marking of a folio or bifolium by a point or knife to guide ruling. - Markierung eines Blattes oder einer Lage mittels kleiner Einstiche als Orientierung für das Liniennetz und den Linierungsrahmen.

0004 Ruling - Linierung

The process by which a frame and/or horizontal lines are produced (via drawing, scratching or folding) to guide the hand in writing. - Prozess der Linierung, Anlegen des Hilfsliniennetzes (Rahmen oder horizontale Schreiblinien) durch Vorzeichnung, Vorritzung oder Knicken/Falten der Seite.

0005 Mise en page / Page design / Layout - Seitengestaltung, Seiteneinrichtung, Seitenlayout

Refers to the layout of a page which means the general arrangement of the various elements on a page. - Bezieht sich auf die Seitengestaltung, die allgemeine Einteilung und Aufteilung der unterschiedlichen Elemente auf der Seite.

0009 Illumination - Buchmalerei, Ausstattung

The embellishment of a manuscript with luminous colours (especially gold and silver), i.e. the total of decorative elements and pictorial representations. - Verschönerung der Handschrift mit leuchtenden Farben (v.a. Gold und Silber); die Gesamtheit der dekorativen Elemente und bildlichen Darstellungen.

0006 Marginalia - Marginalie, Randnotiz

Refers to writing or decoration in the margins of a manuscript, including glosses, reference signs, annotations, and diagrams. - Beschriftung oder Dekor an den Seitenrändern der Handschrift, beinhaltet Glossen, Verweiszeichen, Anmerkungen und graphische Darstellungen.

0008 Script / Lettering - Schrift, Schriftart

The handwriting used in Medieval manuscripts which was subject to more rigid rules and hierarchies than modern personal handwriting (less features of an individual handwriting). - Handschrift eines Schreibers in mittelalterlichen Handschriften, deren Ausführung strikteren Regeln und Hierarchien unterlag als moderne Handschriften (besitzt weniger Merkmale einer individuellen Handschrift).

0007 Initial - Initiale

Enlarged and decorated letter, introducing an important section of a text (work, chapter), emphasised by size, colour form, and decoration. Initials can have different levels of significance (primary, secondary, tertiary). - Schmückender Buchstabe am Textanfang, hervorgehoben durch Größe, Farbe, Form und Verzierung, leitet einen hervorzuhebenden Abschnitt eines Textes (Werk, Kapitel) ein. Initialen haben verschiedene hierarchische Bedeutungsebenen (primär, sekundär, tertiär).

CodiKOS

Browse SKOS Concepts Illustration - Buchmalerei, Ausstattung (0009)

[View CodiKOS on GitHub »](#)

CodiKOS / 0009

0152 Miniature/ Full-page miniature - Miniatur/ Bildseite

Umbrella term for illustration in manuscripts. Refers to an illustration or painting executed in a manuscript, which is independent from borders or initials. The pictorial element may occupy also a whole page. - Überbegriff für Bebilderung in Handschriften. Bezeichnet alle selbständigen, von anderen dekorativen Elementen (Randornamente, Initialen) unabhängigen bildlichen Darstellungen. Auch eine ganze Seite kann mit einer Miniatur ausgefüllt sein.

0153 Painting - Farbe

Colored preparation to cover a certain area. - Farbe zum Auftragen, flüssig oder in Form einer Paste.

0154 Ornamentation / Decoration - Ornamentik, Dekoration

Decoration which does not belong to the illustration itself. - Schmuck / Dekoration, der nicht zu einer Illustration an sich zugehörig ist.

0155 Decorative repertoire - Ornamentrepertoire, dekoratives Repertoire

All patterns and decorative processes employed by an artist, a workshop, a school. - Umfasst alle Muster und Techniken, die einer Malschule, einem Künstler oder einem Atelier zugeschrieben werden können.

0156 Decorative element - Ausstattungselement

Component of the book illumination. - Bestandteil der Buchausstattung.

0157 Infilled ornament - Füllmotiv

A decorative element, that has been inserted subsequently. - Nachträglich eingefügtes Bildelement.

0158 Illustration - Illustration

Pictorial representations of objects, persons and scenes with relation to the content of the text. - Bildliche Darstellungen von Objekten, Personen und Szenen mit Bezug zum Textinhalt.

0159 Cycle of illustrations (see also "Register") - Bildzyklus (siehe auch "Registerbild")

Sequence of representations illustrating the successive phases of a story or the different sections of a text. - Serie von bildlichen Darstellungen, die den Inhalt einer Geschichte wiedergeben und den Text illustrieren.

0160 Decorated, illuminated border / Decorated marginalia - Randillustration, Randminiatur, Kommentarillustration

Decorative element with relation to the text content which is situated on the margin of the page (≠ marginal drawing). - Textbezogene, dekorative Darstellung auf einem der Seitenränder (≠ Randzeichnung).

0161 Panel / Compartment - Bild, Bildfeld

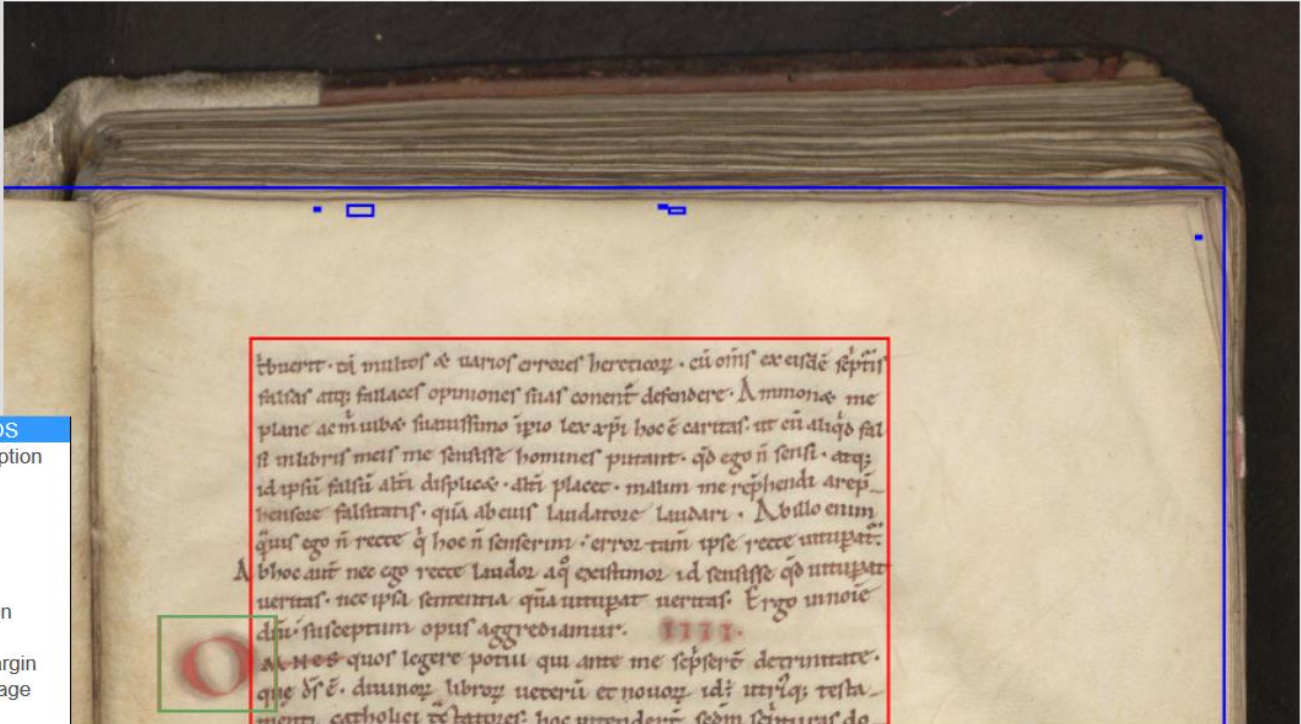
Predefined space which contains figural decoration. - Vorgegebener Raum, der mit einer figurlichen Darstellung ausgefüllt ist.

0162 Bas-de-page(-scene) - Bas-de-page (-Szene)

CodiAnn: Tag

T0152-00011
← Previous
Next →
Save

Pan Verify Modify Delete Tag Draw ▾



Tags Overview

- Illumination
- Body of page
- Initial
- Automatic annotations

Selected Annotation

Initial

[\[http://www.ecodicology.org/CodiKC\]](http://www.ecodicology.org/CodiKC)

[\[http://www.ecodicology.org/CodiKOS\]](http://www.ecodicology.org/CodiKOS) CodiKOS

- [\[http://www.ecodicology.org/A\]](http://www.ecodicology.org/A) Physical description
- [\[http://www.ecodicology.org/B\]](http://www.ecodicology.org/B) Dimensions
- [\[http://www.ecodicology.org/321.01\]](http://www.ecodicology.org/321.01) Pricking
- [\[http://www.ecodicology.org/322.01\]](http://www.ecodicology.org/322.01) Ruling
- [\[http://www.ecodicology.org/331.01\]](http://www.ecodicology.org/331.01) Layout
- [\[http://www.ecodicology.org/331.07\]](http://www.ecodicology.org/331.07) Margin
- [\[http://www.ecodicology.org/511.01\]](http://www.ecodicology.org/511.01) Illumination
- [\[http://www.ecodicology.org/311.05\]](http://www.ecodicology.org/311.05) Page
- [\[http://www.ecodicology.org/331.071\]](http://www.ecodicology.org/331.071) Inner Margin
- [\[http://www.ecodicology.org/331.09\]](http://www.ecodicology.org/331.09) Body of page
- [\[http://www.ecodicology.org/333.02\]](http://www.ecodicology.org/333.02) Initial

tuerit. et multos de uariis erroribus hereticorum. cum omni ex eisdem scriptis
 factis atque fallacis opinionibus suis conent defendere. Ammonens me
 plane ac in uerbis suauissimo ipso. lex aperi hoc est caritas. ut cum aliqdo fal
 si uuldris meis me sensisse homines putant. qd ego n̄ sensi. atq;
 id ipsi falsi alii displice. alii placeat. malum me reprehendi arep
 hentes falsitatis. quia ab eius laudatore laudari. Ab illo enim
 quis ego n̄ recte q̄ hoc n̄ senserim. error tam ipse recte utupat.
 A hoc aut̄ nec ego recte laudo. aq̄ existimor. id sensisse qd utupat
 ueritat. nec ipsa sententia qua utupat ueritat. Ergo innoce
 nti. susceptum opus aggrediamur. **IIII.**
 ad nos quos legere potui qui ante me scripser̄ de trinitate.
 que d̄ e. diuinor. libror. ueteru et nouor. id est utriq; testa
 menti. catholicor. rectoror. hoc uerendū. sedm̄ scimus do

Life Cycle of Annotation Teil 1



Digitalisierung der Handschriften



Speicherung der Digitalisate
mittels Bitstream Preservation



Auszeichnung mit Metadaten
Informationen aus den
Handschriftenkatalogen
(bibliothekarisch, kodikologisch)



**Automatische Extraktion äußerer
Merkmale** auf den Digitalisaten



Anreichern der Metadaten durch
automatisch generierte Tags
(kodikologisch)



Life Cycle of Annotation

Teil 2

SemToNotes

Manuelles Markieren
von Layoutmerkmalen im
Digitalisat (CodiAnn)



Manuelles Tagging durch
Verknüpfung mit
kodikologischer Ontologie
(CodiKos)



crossmediale Verknüpfung
von **Bild und Transkription** (TBLE)

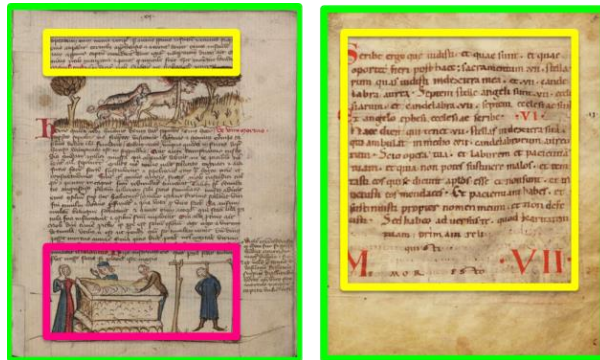
Anreichern der technischen und
bibliographischen Metadaten
(Metadaten Editor)



eCodicology: Was wird annotiert?

Seitengröße
Schriftraum
Bildraum

Höhe, Breite, Anzahl, Position: Messdaten für jede Seite



+



=

```
<tei:TEI>
<tei:teiHeader>
<tei:msDesc>
<tei:layoutDesc>
<tei:layout>
<tei:dimensions type="written" corresp="#written1">
<tei:height quantity="250" unit="mm" min="249" max="251"
confidence="0.8">250mm</tei:height>
<tei:width quantity="100" unit="mm" min="98" max="101"
confidence="0.77">100mm</tei:width>
</tei:dimensions>
<tei:dimensions type="image" facs="#image1"/>
</tei:layout>
<tei:layout columns="1" writtenLines="26 30"/>
</tei:layoutDesc>
</tei:msDesc>
</tei:teiHeader>
</tei:TEI>
```

 page size,  written size,  pictorial space

Woraus bestehen die Annotationen?

```
-<Page_Measure>
  -<zone>
    <zoneID>0</zoneID>
    <CodexID>T Hs 152 1230_4_056_Seite_</CodexID>
    <ImageID>T0152-00056-DEF</ImageID>
    <Area>534.518</Area>
    <Min>0</Min>
    <Max>255</Max>
    <BX>4.288</BX>
    <BY>3.610</BY>
    <Width>20.331</Width>
    <Height>27.508</Height>
    <Major>30.170</Major>
    <Minor>22.558</Minor>
    <Angle>89.866</Angle>
  </zone>
  -<zone>
    <zoneID>1</zoneID>
    <CodexID>T Hs 152 1230_4_056_Seite_</CodexID>
    <ImageID>T0152-00056-DEF</ImageID>
    <Area>0.047</Area>
    <Min>255</Min>
    <Max>255</Max>
    <BX>20.508</BX>
    <BY>3.653</BY>
    <Width>0.424</Width>
    <Height>0.169</Height>
```

Roh-XML für die
Seitenmaße
nach der
Merkmalsextraktion

TEI <msDesc> subset (fiktiv)

```
124 <layoutDesc>
125   <layout>
126     <locusGrp>
127       <locus from="1" to="100" xml:id="locusGrp002"/>
128     </locusGrp>
129     <dimensions type="written" corresp="#locusGrp002">
130       <height quantity="250" unit="mm" min="249" max="251" confidence="0.8">250 mm</height>
131       <width quantity="100" unit="mm" min="98" max="101" confidence="0.75">150 mm</width>
132     </dimensions>
133     <dimensions type="image">
134       <height quantity="50" unit="mm">50 mm</height>
135       <width quantity="50" unit="mm">50 mm</width>
136     </dimensions>
137     <dimensions type="image" facs="#zone01"> </dimensions>
138   </layout>
139   <layout columns="1" writtenLines="26 30"/>
140 </layoutDesc>
141 </objectDesc>
142 <decoDesc>
143   <decoNote type="miniature" facs="#zone01">
144     <p>One full-page miniature, facing the beginning of the text.</p>
145   </decoNote>
146 </decoDesc>
147 <additions>
148   <note type="gloss" place="margin"></note>
149 </additions>
359 <facsimile>
360   <surfaceGrp type="opening">
361     <surface type="leaf" ulx="0" uly="0" lrx="200" lry="300">
362       <graphic url="T0099-00005.tif"/>
363       <zone type="image" ulx="20" uly="20" lrx="70" lry="70" xml:id="zone01"/>
364     </surface>
365   </surfaceGrp>
366 </facsimile>
```

SVG in XML

```
<facsimile>
  <!-- metadata describing of the digital images -->
  <surfaceGrp type="opening">
    <surface type="leaf" ulx="0" uly="0" lrx="200" lry="300">
      <graphic url="0099-00005.jpg"/>
      <!-- Maße der Seite -->
      <zone type="image" ulx="20" uly="20" lrx="70" lry="70"
xml:id="image1"
      ><!-- Schrift- und Bildfelder in Koordinaten --></zone>
    </surface>
  </surfaceGrp>
</facsimile>
```

Wozu dienen die Annotationen?

Metadaten aus dem Virtuellen Skriptorium St. Matthias (liegen in METS-Dateien mit integriertem TEI Header vor):

- **technisch**
- **administrativ** (bez. auf gesamten Kodex)

Auskunft über Besitz und Provenienz der digitalen Ressource

- **deskriptiv** (bez. auf gesamten Kodex)

Angaben zur Quelle, wissenschaftliche Beschreibung, z.B. Datierung, Beschreibstoff, Buchformat, Blattzahl, Inhalt; rudimentäre kodikologische Angaben

Kombination mit den automatisch annotierten Messergebnissen aus eCodicology, z.B. Seitenhöhe, Seitenbreite (Seitenfläche), Anzahl der Textspalten, Anzahl unbeschrifteter Seiten usw.

→ **Statistische Analysen**

Seitengröße vs. Beschreibstoff

Stichprobe 50	Pergament	Papier
Mittelwert Seitengröße (cm ²)	349.94	442.91
Median Seitengröße (cm ²)	350.55	442.78
Stichprobe 100	Pergament	Papier
Mittelwert Seitengröße (cm ²)	339.06	390.72
Median Seitengröße (cm ²)	339.34	391.45

Seitengröße vs. Spaltenzahl

Stichprobe 50		
Seitenfläche	einspaltig	zweispaltig
MITTELWERT (cm ²)	352.17	396.19
MEDIAN (cm ²)	353.01	396.18

Stichprobe 100		
Seitenfläche	einspaltig	zweispaltig
MITTELWERT (cm ²)	328.29	456.91
MEDIAN (cm ²)	353.98	456.43

Geplantes Web Portal

Ziele

- 1) Einbringen von fremden Handschriftendigitalisaten, an denen die Algorithmen trainiert werden können
- 2) Möglichkeit der Abfrage, z.B.:
 - Ausgehend von bestimmten Kodizes, z.B. „Zeige mir alle zweispaltigen Sequenzen in diesem Kodex/in diesen Kodizes an“ etc.
 - Ausgehend vom Gesamtbestand, z.B. „Suche mir alle Kodizes aus dem Zeitraum X, in denen das Verhältnis von Seitenfläche und Textfläche Y:Z beträgt“ etc.
- 3) Visualisierungsmöglichkeiten