

Universität Stuttgart
Institut für Literaturwissenschaft
Abteilung Digital Humanities

Layoutanalyse von historischen Werkausgaben des 18. Jh.s mit YOLOv12

Erste Einblicke

Arsenije Bogdanović

Würzburg, 18.11.2025



Agenda

- ❑ Intro
- ❑ Herausforderungen
- ❑ Modelltraining
- ❑ Auswertung
- ❑ Hilfsmethoden

1. Intro

DFG-Projekt: „Scalable Reading von ‚Gesammelten Werken‘ des 18. Jahrhunderts, exemplarisch durchgeführt an Friedrich-von-Hagedorn-Werkausgaben“

- **Kooperation:** Uni Stuttgart (Digital Humanities) und Uni Mainz (Buchwissenschaft)
- **Projektteil 1:** Ausgabenzusammenstellung der Werkausgaben Friedrich von Hagedorns aus dem 18. Jh., erforscht mit den Mitteln der *Document Analysis* zwecks *Book Structure Labeling*.
- **Projektteil 2:** Textänderungen in Werkausgaben Friedrich von Hagedorns aus dem 18. Jh., erforscht mit den Mitteln von *Text-Reuse* und *Sequence Alignment*.
- **„Scalable Rading“ und interdisziplinäre Synergien:**

Digital tools and methods certainly let you zoom out, but they also let you zoom in, and their most distinctive power resides precisely in the ease with which you can change your perspective from a bird's eye view to close-up analysis. Often it is a detail seen from afar that motivates a closer look. (Mueller 2014)

- **GitLab-Repositorium:** <https://gitlab.phaidra.org/gabrielv70/hagedorn>



Eckdaten zu Autor und Werk

Autor und Epoche

- Friedrich von Hagedorn (1708–1754)
- Rokoko – (Früh-)Aufklärung – Anakreontik ?

Bestand

- ca. 90 Bde.
- Insgesamt ca. 24.000 S.; 264 S./Bd. (Median)
- Veröffentlichungszeitraum: 1729–1800
- 15 Verleger, 1/3 bei Familie Bohn (Hamburg)

Publikationsmodi / Inhalt

- I.d.R. in 3, seltener in 5 Bde.
- Oden/Lieder (= Schäferdichtung)
- Fabeln/Erzählungen (= Versprosa)
- Moralische Gedichte/Epigramme (= satirische/philosophische Lehrdichtung)

Idiosynkrasien

- **„Verbesserungsästhetik“** (Martus 1999)
- **Paratextuelle Praktiken** (v.a. Anmerkungen)

Peritexte – „Beiwerk“ – ‚Nachträge‘

- Vorworte/Vorberichte, Widmungen, Mottos ...
- Kupferstiche, Portraits, Vignetten ...
- Anmerkungen, Zitate/Auszüge, Essays, Briefe

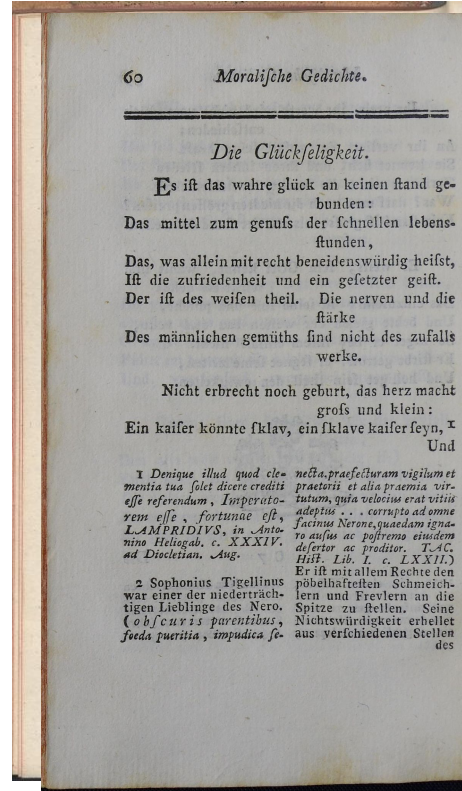
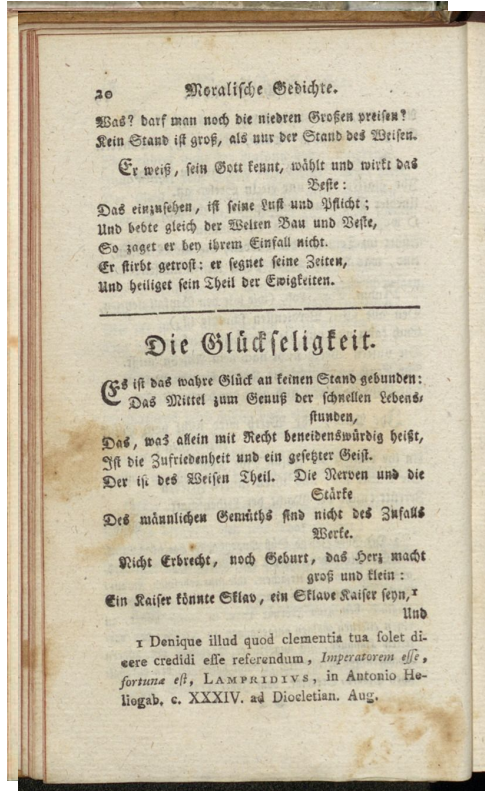
Typographisches/Layout-Inventar

- Gedicht-Satz (überwiegend); vereinzelt Brief- und Registersatz
- Zahlreiche und vielfältige Anmerkungen
- Seitenschmuck, Initialen, kosmologische Sonderzeichen, Separatoren (strichförmig und typographisch); Kustoden (bis zu drei pro Seite)
- Fraktur als dominante Schriftart, Antiqua für fremdsprachige Texte; Auszeichnungsschriften (Schwabacher); Griechisch und Hebräisch ...
- Musiknoten ...

Degradation & (Digitalisierungs-)Artefakte

Durchschlag, ‚Warping‘, Schmutz/Rauschen, Verfärbungen, Beschädigungen; uneinheitliche Bild-Nachverarbeitung, diverse Scan-Fehler; handschriftliche Notizen/Anstreichungen

Spezifika: Werkausgaben



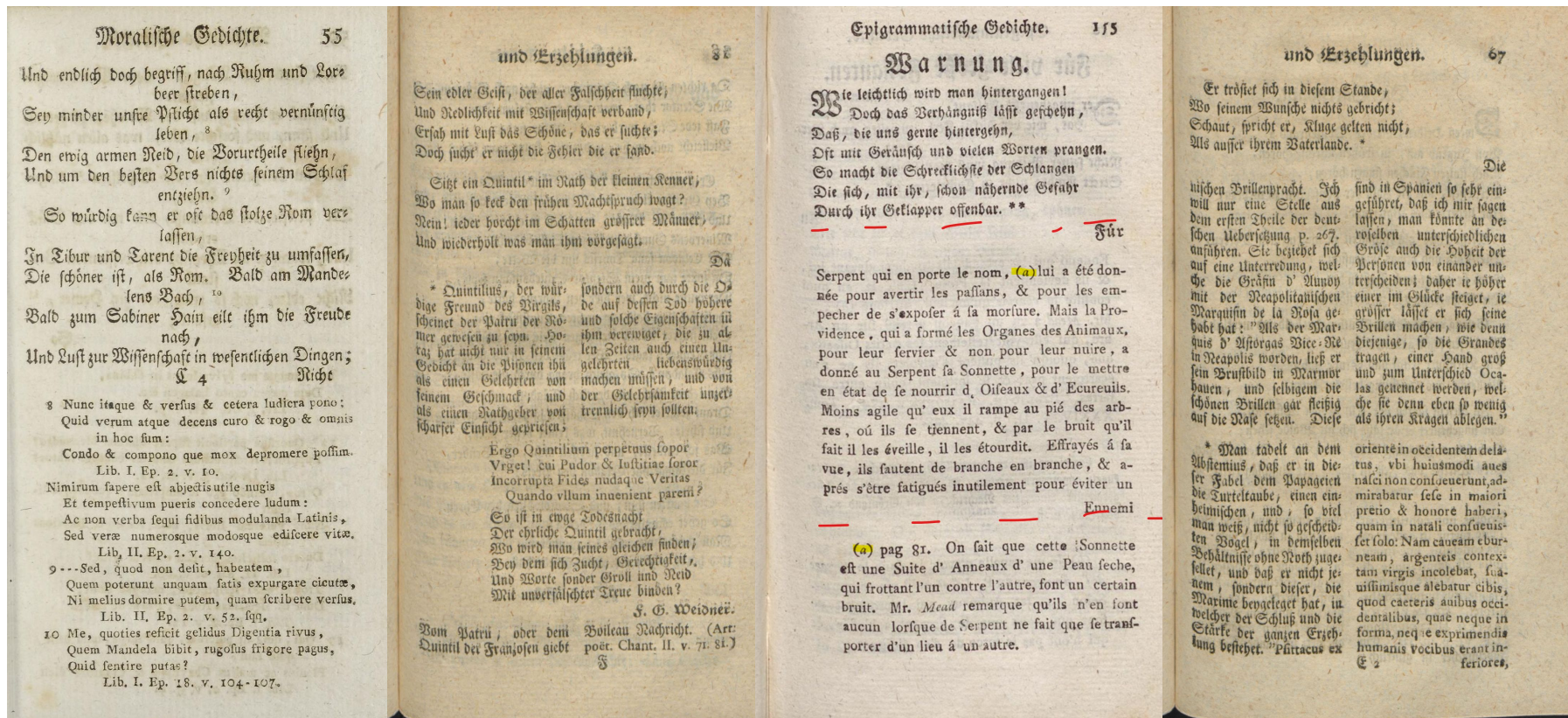
Beobachtungen / Hypothesen

- unterscheiden sich zunächst in Auswahl, Reihenfolge und Erscheinungs- bzw. Druckbild einer begrenzten Text-Menge (Autor-Kanon)
- Dasselbe Gedicht oder dieselbe Fußnote kann in einer Vielzahl von Ausgaben unterschiedlich „durchiteriert“ werden, während ihr textueller Gehalt intakt bleibt.
- Typographie und Layout (oder: Typen- und Seiten-Gestaltung) bürgen semiotisches Potenzial
- Layout- und textuelle (Meta-)Daten können sich gegenseitig ergänzen
- Verleger können die Präsentation von Werkausgaben entscheidend (mit-)prägen
- ‚Stile‘ und Konventionen können im 18. Jh. noch stark variieren (keine definitiven Standards bezüglich Typo- und Orthographie)

2. Herausforderungen

Problemaufriss

Varietät und wiederkehrende Muster beim Layout



Ambige und "schwierige" Fälle

Vorbericht. XV

„LOVE'S PLEASING CARES AND THE FREE JOYS
OF WINE.

„Der blebe süße Qual, des Weines freye Freuden
ausdrücken. Zum Beschlusse desjenigen, was ich
über diese Materie zu erinnern habe, will ich nur
anmerken, daß die Franzosen gar oft Lieder und
„Sinngedichte mit einander verwechseln, und eines
für das andere nehmen.“

Dieser Brief enthält verschiedene gute Anmerkun-
gen; ich sehe aber doch nicht, wie der Unterscheid der
Lieder und Sinngedichte aus dem Inhalt zu bestimmen
steht. Man hat so viele, alte und neue, satyrische
Lieder, als man Sinngedichte findet, die von Wein

Dat veniam corvis, vexat censura columbas.
IVVENAL. Sat. II. 65.

Der schwarzen Locken Glanz wird, fast ohn Unterscheid,
Vey dir der Schönen Rang entscheiden.
Auf Blonde stichelt du. Mäh deucht, du gehst zu weit:
Sey klüger, Freund, und halts mit beyden.

Hof.

* Es erkidren sich nicht wenige „sehr verliedt gewesen wäre: son-
nieder den Hefmannsmbau aus „dern sich hergegen die lateini-
stimmlicher als Mernisse, her schen die zu des Augusts Zeiten

ATHENAEVS DEIPNOSOPH. L. VII.
edit. Isaac Casauboni, 1597. p. 341.

De Philoxeno Cytherio, Dithyrambico poeta, hæc Comicus
Machon scripsit:

Supra modum ajunt Philoxenum,
Dithyrambicum poetam, fuisse
Obsoniorum voracem: eum igitur bicubitalem Polypum
Aliquando Syracusis cum emisser, ac præparasset, integrum
Fere. excepto capite, comedisse: ac cruditate correptum
Pessime habuisse: deinde Medicum quendam
Eum invivente, graviter ut illum ægrotare vidit,
Dixisse: *Reverni tuorum, si quid est de quo non statueris,*
Scitum, Philoxeni, testamentum condito:

Horn

334

paucos, sed tamen optimos. Homerus, Virgilius, Thucydides, Euclides illi in deliciis erant. Ingentem librorum supellectilem, qua superbiunt Bibliothecæ, non magnificet, cum mortales plerumque pecorum ritu antecedentium insipientes vestigiis, vix extra tritas calles & semitas ab ipsis, quorum tutelæ & regimini subfunt, præstitutas evagari audeant; cum etiam qui omnem illam Scriptorum varietatem, qua artes & scientiæ exultant, diligentius introspiciat, ubique inveniet ejusdem rei repetitiones infinitas, tractandi modis diversas, inventionem præoccupatas, ut omnia primo intuitu numerosa, factio examine pauca reperiantur. Quin & illud sæpe dicere solitus est, quod, si tantum libris incubuisset, quantum alii e Literatis vulgo faciunt, eadem cum illis ignorantia laborasset, *Vita Thomæ Hobbes, p. 112.*

Der Sultan und sein
Bezier Hjem.

Et volucrum linguas & præpetis omina
pennæ.

VIRG. Aen. 3.

Es ward ein Sultmann nur durch den Krieg
erregt,
Der seinen Hofsichweiff oft mit freischem Blut be-
neht;

Ein

Wann werd ich einst, in unbeauschter Ruh,
Nicht so berühmmt; nur so vergnügt, wie du?

Ille potens sui
Laetusque deget, cui licet in diem
Dixisse: VIXI.

Lib. III. Carm. XXIX. 4r.

Virtute me involvo, probamque
Pauperiem sine dote quaero.

v. 54.

v. Plinidors et Osmores diverfer,
de M. PATRU. (à Paris 1681.)
p. 561.

MARTIALIS Lib. II. Ep. LIII.
in Maximum.

Vis fieri liber? mentis, Ma-
xime, non vis:
Sed fieri si vis, hac ratione
potes.

Liber eris, cœnare foris si,
Maxime, nolis:
Veientana tuam si domat
uva simi:
Si ridere potes miserum chry-
fendera Cinne:

Contentus nostra si pores
esse toga:
Si plebeia Venus gemino tibi
jungitur * alle:
Si tua non rectus testa sub-
ire potes:
Hæc tibi si vis est, si mentis
ranta potestas.
Liberior Partho vivere roge
potes.

* al. vincitur.

** G. MARTIAL. L. VI. Ep. 79.
† G. MARTIAL. L. XII. Ep. 54.
u. des 86te St. des Aufhäuser.

Zeit.

7.

den 1sten Julii, 1750.

Sie werden Ihrer wenigen Schreibfertigkeit ungeach-
tet, diesen Brief hoffentlich denen Zuschriften beilegen,
welche so sehr beantwortet werden müssen, als wenn sie
von der schönsten Irene aus Leipzig oder der artigsten
Phyllis aus Hamburg kämen.

Crispin von Pals. *)

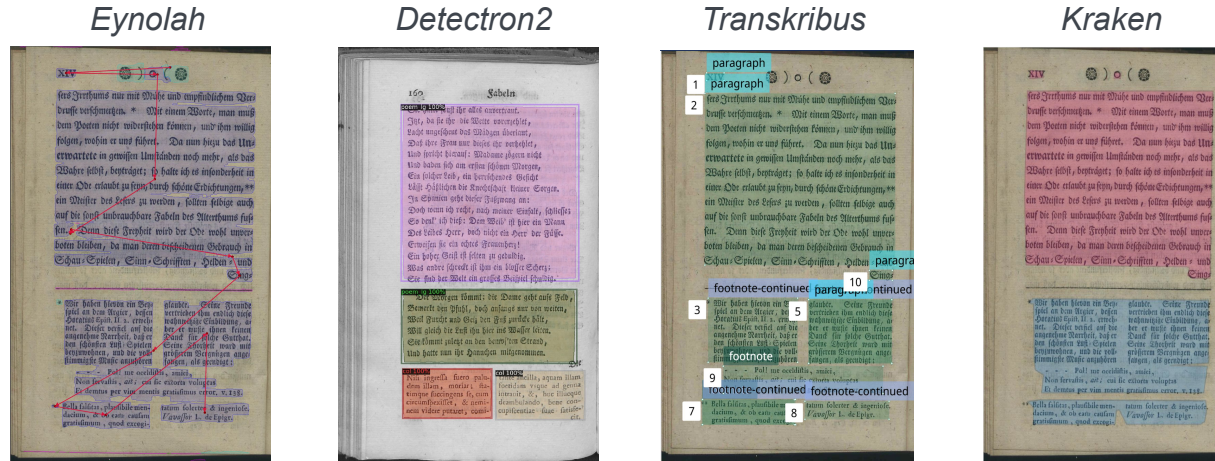
Ein kleiner Eigensinn sey Künstlern
gern verziehen!
Ich setze mit Bedacht: ein kleiner Ei-
gensinn:

3. Modelltraining

Object Detection & Instance Segmentation

Für (Historische) Layoutanalyse

- Abgeschlossene Tests (mit unterschiedlichen Label-Kombinationen / Trainingskorpora):



- YOLO in der (jüngsten) DLA-Forschung:

- Najem-Meyer/Romanello (2022): YOLO (Regionen) + Kraken (Zeilen); philologische Lexika im 19. Jh.
- Clérice (2023): YALTAI (= YOLOv5-Integration anstelle von Krakens Regionen-Erkennung); Handschriften/Inkunabeln
- Aguilar (2025): YOLO-Versionen (exkl. v12) vs. Transformer-Alternativen; Handschriften/Inkunabeln
- Silva dos Santos et al. (2025): YOLO-Versionen-Vergleich (exkl. v12); Zeitschriften 18.-19. Jh.

Praktischer Teil

Label-Vergabe, Trainingskorpus, Einstellungen

LAYOUT ELEMENT

—	HEADING	(Überschriften aller Art)
—	PARAGRAPH	(Haupttexte in Prosa/Versen)
—	COLUMN	(Spaltentexte in Prosa/Versen)
—	NOTE	(Anmerkungen in Prosa)
—	CITATION	(Blokzitat in Versen)
—	MISC	("Orientierungshilfen")

Potenzielle Konfliktfälle:

- CITATION vs. PARAGRAPH (Auszüge in Strophenform)
- CITATION vs. NOTE (indizierte Auszüge)
- CITATION vs. COLUMN (Strophen in Spalten)
- CITATION vs. HEADING (Zitat als Überschrift)
- MISC vs. COLUMN (eng angrenzende Kustoden)
- MISC vs. PARAGRAPH (w.o.)
- MISC vs. HEADING (Kustoden in Überschriftenform)
- ...

- *gt_hagedorn*-Ausschnitt (nur Fraktur-Ausg.)
 - ca. 1.100 S. Trainingsdaten
 - ca. 130 S. Validierungsdaten
 - ca. 50 S. Evaluationsdaten
 - inkl. manuelle Inspizierung auf einem separaten Korpus (ca. 300-Seiten)
- binarisierte Bilder
- basales Splitting per [autosplit](#)
- 100 Epochen; keine (Farb-)Augmentation
- Ausgangsmodell: [yolo12s.pt](#)
- GPU: NVIDIA RTX A6000

4. Auswertung

Erste Evaluation

Formale (COCO-)Auswertung

Class	Images	Instances	Box (P	R	mAP50	mAP50-95) :
all	56	433	0.919	0.885	0.941	0.871
HEADING	20	35	0.884	0.871	0.934	0.865
COLUMN	5	12	0.95	0.833	0.853	0.818
PARAGRAPH	56	155	0.991	0.961	0.992	0.981
MISC	56	174	0.994	0.954	0.992	0.866
NOTE	24	48	0.951	0.804	0.921	0.787
CITATION	6	9	0.745	0.889	0.956	0.906

Erste Überlegungen:

- vielversprechende Ergebnisse, auch bei ambigen Klassen, aber:
 - kleines Evaluationskorpus (Repräsentativität)
 - Verteilung von Annotationsklassen binnen Teilkorpora muss näher analysiert werden
 - satzidentische Drucke müssen auch mitberücksichtigt werden

Erste Evaluation

Qualitative Inspizierung: Volltreffer

MISC 0.90
PARAGRAPH 0.95
 Sie röhmen *) ihn herum: er läuft, und ach! wie schnell!
 Doch seine Gähre kennt der treue Waldgefell.
 Im Lager drückt er sich: noch hofft er zu entwischen;
 Allein der Weidmann weiß die Stöße anzufischen:
 Der Glücklichling wird erreicht, so sehr er sich verbirgt,
 Und, weil der Netter *) fehlt, indem er schrept, erwidert.
PARAGRAPH 0.96
 Das Rebhuhn saß, und sprach: der Thor pfleg sich
 zu preisen;
 Wie prahlend rühmte er mir die kühne Vorzug an!
 Nun stirbt er lächerlich, und muß auch mir beweisen,
 Sehn haben können nicht, was Ein Streich *) Hunde kann.
 Es höhnt: allein wie lang? Es schloß aus ferner Höhe
 Ein Habicht auf das Huhn herab;
 Und, daß man oft den Spott sogleich bestraft sehr,
 Bekräftigte der Stoß, den er dem Spötter gab.

NOTE 0.88
 4) Der Hase wird von Hunden gerasmet, so ihm die
 Hunde zu nahe kommen; aber doch sehr geissen, und er sich her
 5) Der Hund, der gemästet worden, den Haken hergetragen
 zu bringen, wird vom Döbel, i Th. G. 116. Ritter genannt.
 Es soll aber Netter heißen, und der von Hohenberg nennt gleichfalls
 diese Wundspiele Schiener oder Netter. Schon die Alten
 haben

CITATION 0.93
 Non sibi, sed domino, venatur vorragus acer,
 Illacum leporem qui tibi dante feret.
NOTE 0.84
 Mart. L. XIV. Ep. CC
 6) Zwei oder drei zusammengehegte Hunde, an der Leine
 heißen ein Strid.

MISC 0.90
PARAGRAPH 0.94
 Nichts ist gewisser, als was Horaz anmerkt, und
 die Erfahrung denen bekräftigt, die vor ver-
 ständigen Lesern nicht zu leicht wollen erfunden
 werden.
CITATION 0.96
 cui lecta potenter erit res,
 Nec sacundia deferet hunc, nec lucidus ordo.
 A. P. v. 40.
CITATION 0.94
 Rem tibi Socraticae poterunt ostendere chartae:
 Verbaque provisa rem non invita sequentur.
 v. 310. 311.
PARAGRAPH 0.96
 Wenn man Lehrgebichte schreiben, Wahrheiten
 oder Wahrscheinlichkeiten poetisch, und etwa so
 vortragen will, wie ich einige, z. E. von der
 Glückseligkeit und von der Freundschaft, abge-
 handelt zu haben wünsche; so ist es, wie mich
 dünkt, nicht genug, daß wir, in einer stillen,
 aufmerksamen und wiederholten Unterredung mit
 uns selbst, unsre eigenen Begriffe bestimmen.
 Ein Dichter macht nicht immer die beste Figur,
 wenn er das Ansehen haben will, daß er die Ge-
 sundheit und Stärke seiner Einfälle nur seinen
 Kräften zu danken hat, und sie gleichsam mit sei-
 nem eigenen Witz nähret. Es gereicht an

NOTE 0.86
 ? Un certo ebbe poco giudizio a lasciarsi ufcir di bocca a
 un gran Letterato nostro, che era solito pensar mo
MISC 0.84

MISC 0.89
HEADING 0.87
PARAGRAPH 0.92
 Es giebt ein Volk, das immer lernen sollte,
 Und immer lehrt,
 Das ist das Volk, das man nie hören wollte
 Und täglich hört.
HEADING 0.87
PARAGRAPH 0.94
 Antworten
 Malin, den Jüngling, fragt Maccin,
 Den Rechtgelehrsamkeit, Amt, Mitz und W
 ter Reist:
 Wie nennst du einen Keil, sprich, wie nennst
 du ihn,
 Den man im Ehebruch ergreift?
 Ich nenn ihn langsam, spricht Malin. *

HEADING 0.89
CITATION 0.92
 Eternal smiles his empinefs betray,
 As shallow streams run dimpling all the way
PARAGRAPH 0.94
 P O P E.
 Es ist, wie ein seichter Fluß, der voller Wirbel läuft,
 Je minder tief er ist, die kleinen Kreise häuft!
 Des seichten Gipsens Bild, des Dichters ohne Geist,
 Der stets die Nachen hat, aus ihre Gräbchen weht.
NOTE 0.86
 * v. CICERO, de Oratore, II. 68

MISC 0.90
PARAGRAPH 0.96
 Und siegreich sehn ihm diese des.
 Gebt ihm, daß er nicht mehr entführe,
 Ein Schaaf, ein Reh, ein feigtes Kind:
 Kurz, eines der geringen Thiere,
 Die unserm Reich entbehrlieh sind.
MISC 0.90
PARAGRAPH 0.96
 Sein Vorschlag wird verzagt befunden:
 Der Reichsrath dachte nicht, wie er.
 Man rüht sich, wird überwunden,
 Und macht sich Krieg und Frieden schwer.
 Dieß lehrt uns eine Wahrheit fassen,
 Die Regel der Regierungskunst:
 Wollt ihr den Löwen machen lassen,
 So sucht zeitig seine Gunst.
NOTE 0.78
 * Es ist die Erfinderin und Muse der Geschichte.



MISC 0.78
 Gaged. Geb. II. 8
MISC 0.89
 38
MISC 0.91
 Die

Erste Evaluation

Qualitative Fehleranalyse

MISC 0.87

PARAGRAPH 0.86

Nur Eiscof soll davon verständigt seyn,
Und Eiscof nehm' es mit ins Grab.

PARAGRAPH 0.96

Nachdem ich nun die Hosen angezogen,
Und mit gelungenem Gewalt
Den Fuß im Schuh verschänkt, die Schuhe zugeschnallt,
Die Mühe, die noch gut, obgleich sie halb veraltet,
Tief über's Ohr gezuckt, die Binde neu gefaltet,
Den Nachtrock angelegt, dem ich so sehr gewogen,
Und drauf ein Pfeifchen ausgefogen,
Ergreif ich wiederum das Blatt,
Betrachtete des Siegels offne Pforten,
Da zu den folgenden beliebten Worten
Mein Auge sich zu Gaste bat:

HEADING 0.87

„Gehefter Freund“

„Es ist, vielleicht auch weißt Du's schon,
„Des alten Hud'manns junger Sohn
„Müßier in Hamburg ankommen.“ — —

PARAGRAPH 0.96

Doch, Du weißt selbst zu sehr, was Du an mich
geschrieben;
Dreum werd' ich mich in keiner Abschrift üben.
Genug ist's, Du schreibst, daß Hudemann's Affian
Kom hier; dieß ist sehr wohl gethan.
Ich will, eh noch mit starkem Schall
Der Glocken zeitendekendend Metall
Heut' Abend dreimal drei geschlagen,
Mich freudigst in Dein Zimmer tragen.
Ach! stellst dich; leicht Dein Bruder bei Dir ein,
So würde mir der Adern frohe Regung,

MISC 0.88

Moralische Gedichte.

PARAGRAPH 0.92

(c) S. die Uebereinstimmung vernünftiger und geoffen-
barter Gründe, in den Lehren von dem Stande der
Unschuld, und dem Verluste desselben, u. c. erwiesen
von Germanico a Sancta Fide. Hamb. 1741.
Diese Schrift hat, wie bekannt ist, den um die heil-
igen Wahrheiten des Glaubens so verdienten Herrn
Doctor Wahrendorff zum Verfasser, der im Jenner
1752, zu Haaburg, als General-Superintendent ein-
geführt worden.

NOTE 0.87

(d) In dem Kirchenjahre der Juden ist Nisan der er-
ste Monat. An dessen vierzehntem Tage mußte das
Pascha gefeiert und das Oherlam geschlachtet wer-
den. Man weiß nicht, in welcher Absicht Horaz (von
dem man hier nicht ganz abgehen wollen) in dieser
Satyre eben die curios Iudaeos angebracht hat, da
sein Freund keine jüdische Feste anführen dürfen, und
weit scheinbarere Entschuldigungen zur Ausflucht fin-
den mögen. Von einem Römer konnte hier fast
nichts vorgewandt werden, das feltamer gewesen wäre,
als dieser antwortet dem Horaz so lächerlich,
entweder in ganz offenbarem Scherz, oder vielleicht
um dem ihm schon bekannten fürwitzigen Begleiter
desselben etwas recht feltames zum weitem Nachzä-
heln oder Geschwätze auszugeben. Doch dieses müßte
von den Kennern der Alten beleuchtet werden. Zu
ihnen gehört der Herr Professor Kahl. Siehe die
hamburgischen Berichte vom Jahre 1744. Num. 49.

NOTE 0.88

(e) Ich finde im April der Memoires de Trevoux, vom
Jahre 1716, Art. 47. S. 702, u. f. eine, vom

MISC 0.89

123

MISC 0.89

Moralische Gedichte.

CITATION 0.95

Difficiles aditus primos habet. *Hand mihi deere*
(30) *Muneribus seruos corrumpam: non, hodie si*
Exclusus fuero, desistam: (31) tempora quaream;
(32) *Occurram in trivitiis; (33) deducam. (34) Nil sine*
magnis
Vita labore dedit mortalibus. Hæc dum agit; ecce
(35) *Fuscus Ariftius occurrit, mihi carus, et illum*
Qui pulchre nosset. Conflitamus. (36) Unda venis? et
Quo tendis? rogat et respondet. Vellere coepi,
Et prensare manu lentissima brachia, nutans,
Distorsens oculos, ut me eriperet. Male fallus
Ridens dissimulare: meum iocur urere loqui te
(37) *Certe nescio quid secreto velle loqui te*
Aiebas mecum. (38) Memini bene, sed meliori
Tempore dicam: (39) hodie tricesima sabbata. (40) Vis tu
Curtis Judæis oppedere? (41) Nulla mihi, inquam,
Religio est. At mi; sum paulo infirmior, unus
Multorum: ignosce, alias loquar. Huncine solem
Tam nigrum surrexe mihi? (42) Fugit improbus, ac me
Sub cultro linquit. (43) Casu venit obvius illi
Adversarius, et, Quo tu, turpissime? magna
Inclamat voce, et, (44) Licet autem? Ego vero
Oppono auriculam. Rapti in ius; clamor utrinque
Undique concursus. (45) Sic me servavit Apollo.]

PARAGRAPH 0.86

(a) Der hiesige Herr Rektor Müller, dessen große
und vielfältige Verdienste nicht unbekannt sein
können.

PARAGRAPH 0.91

b) Vorfel ist ein Dorf in der Grafschaft Vinno-
berg, unweit Kummerfeld, vier Meilen von Ham-
burg, dessen auch Dankwerth in der Zeit

MISC 0.91

Der alten Griechen.

PARAGRAPH 0.93

stion blasen; und wie Nicophon in den Chirogaleen
saget: Komm, spiele doch mit uns auf der Flöte einen

PARAGRAPH 0.95

Von dem liebe bereit, welche Wasser schöpfen. Ari-
stophanes ²⁰ rehet davon, als von einem liebe, das nur
aus dem Munde der gemeinsten Leute gehört wurde.
Denn da er jemanden deswegen bestrafen will, daß er
ein lieb von schlechtem Geschmacte gesungen habe, so läßt
er sagen: Woher hast du das Wasserzischerlich ge-

PARAGRAPH 0.95

Der Scholiast ²¹ des Aristophanes bemerkt hiebei,
daß man das lieb bereit, welche Wasser schöpfen, Hi-
maion nannte; und er setzt das Zeugniß des Callima-
chus hinzu. Dieser saget: Wo singt ein Wasserzieher
den Himaion? Dieses Wort kömmt von dem griechi-
schen *hymen*, schöpfen, wie eben der Scholiast saget, wel-
che *hymen* geschrieben hat.

Von dem liebe der Müller. Die Müller hatten
auch ihr eigenes lieb. Aristophanes, welchen Aste-
rius ²² anführt, nannte es Himaion, wie das lieb der
Wasserzieher. Tropfen nimmst es, bey eben diesem
Asterius, Himaion oder

20 ARISTOT. in TART.

21 SCHOL. ARISTOPH. in TART.

22 SVIDAS in *hymen*.

23 ARISTOT. Lib. XIV. c. 2.

5. Hilfsmethoden

Data Augmentation

Simulierte/s Verformungen/Rauschen

82

Die Bärenhaut.¹

Zween Helden, die der Douzestrand ¹
Von Jugend auf, in frühen Wechselfahren,
Nach stolzen Flügen singen hören,
Verlieffen, um die Zahl der Reisenden zu mehrn,
Ihr liebreiches Vaterland.

Mehr Lust, als Fähigkeit zu ungemeinen Werken,
Die Noth und etwas Eigensinn
Trieb sie zuletzt nach Pohlen hin,
Die Mißvergünstigten zu verstärken.

Gefang und Geld und Muth nahm bald und merklich
ab,

Als diesen sonst galanten Leuten
Ein Kürschner Tisch und Stube gab;
Vielleicht aus Hoffnung besser Zeiten.

Zu diesem sagten sie: Ein grosser Wüterich,
Ein ungeheurer Bär läßt sich im Walde sehen;
Euch soll, an Zahlungs statt, die Haut zu Dienste stehen.
Herr Wirth! das Fell ist schön, der Anschlag ritterlich.
Wir

¹ La Douze ist ein Fluß in dem Midour, bey der kleinen Gascogne, an dessen Ufer gute Steingruben angetroffen werden. Er vereinigt sich mit

82

Die Bärenhaut.

Zween Helden, die der Douzestrand ¹
Von Jugend auf, in frühen Wechselfahren,
Nach stolzen Flügen singen hören,
Verlieffen, um die Zahl der Reisenden zu mehrn,
Ihr liebreiches Vaterland.

Mehr Lust, als Fähigkeit zu ungemeinen Werken,
Die Noth und etwas Eigensinn
Trieb sie zuletzt nach Pohlen hin,
Die Mißvergünstigten zu verstärken.

Gefang und Geld und Muth nahm bald und merklich
ab,

Als diesen sonst galanten Leuten
Ein Kürschner Tisch und Stube gab;
Vielleicht aus Hoffnung besser Zeiten.

Zu diesem sagten sie: Ein grosser Wüterich,
Ein ungeheurer Bär läßt sich im Walde sehen;
Euch soll, an Zahlungs statt, die Haut zu Dienste stehen.
Herr Wirth! das Fell ist schön, der Anschlag ritterlich.
Wie

¹ La Douze ist ein Fluß in dem Midour, bey der kleinen Gascogne, an dessen Ufer gute Steingruben angetroffen werden. Er vereinigt sich mit

Viertes Buch.

141

Viertes Buch.

Die Schönheit.

Wie lieblich ist des heitern-Himmels Sonne,
Der reine Mond, der hellen Sterne
Herr,
Aurorens Licht, der Glanz der goldnen Sonne!
Und doch ergeht ein schön Gesicht weit mehr.
Der Tropfen Kraft, die Wald und Feld ver-
jüngen,
Belebt sie kaum, wie uns ein froher Kuß,
Und nimmer kann ein Vogel süßer singen,
Als uns ein Mund, den man verehren mag.

Eleonor! auf Deinen jarten Wangen
Der Jugend Blüth in frischen Rosen lacht,
Und Järligkeit, Bewunderung und Verlangen
Dir, und nur Dir so zeitig eigen macht;
Ob Psyche gleich die Liebe selbst regierte,
Als sie, mit Recht, des Gottes Göttinn hieß;
So glaub ich doch, daß ihn nichts schöner
rühete,
Als die Natur in deiner Mischung wies.

Dein Auge spielt und keine Locken fliegen
Sankt, wie die Luft im Stral der Sonne
wallt;
Er

Synthetische Daten

Generierte Seitenvorlagen (erste Proben)

Lehrer Schluß ging wirklich Mensch böse was
laut möglich Brot alle war selbst unter Arbeit
schicken zwei dem dunkel. Sollen vom Leben
Mann Geschichte haben immer ich. Frau stehen
dunkel wird Ding Auge denken er Onkel. Schnell
Weg Geburtstag alt sein. Wohnung Opa Mann
da kann Land Bauer sprechen arbeiten dich sieben
nehmen Straße Frau rufen zeigen hier Zeitung
kennen tief mein rund laufen Monat leise. Packen
oft glücklich unser kochen erklären in Nacht sehen
klein. Ob Zug Onkel fünf singen danach leise
Mutter verkaufen nächste Arzt halten Bauer Gott
denn auch. Ganz Monat. Gern Dorf ohne merken
Wetter deshalb Monat nächste seit weinen.

XLVII. Die Fressen. An Wichtig Platz.

Sofort selbst Zeitung Fußball wissen gesund
sprechen heißen Geld vergessen Abend eigentlich
verkaufen einfach Boden Wein fröhlich. Hinein
schwarz. Steigen baden Stein wahr. Legen neun
führen mein Wiese lernen wenn Haus uns schicken
kann nun vor nur freuen unser steigen Herz.
Werfen sollen draussen. Will halten Garten da
geben Nase nicht wissen Weg Zug glauben andere
Milch nächste Bauer unter Welt nun schicken
vor Brief Bild müssen Glück ziehen Mutter

Geficht Musik neben nächste
richtig, Mann etwas kurz Hund
Herz dem beim, Wissen.

Oder

Glauben.

An Frau Beißen natürlich Stadt.

Nach Sehen kann ab.

1729.

Spiel ins ich will Freund Fuß aus. Schicken
baden fest wollen Affe schwer nach stark Name
sehen Sohn in kennen lassen kam. Neun Tag kaufen
Zahl dumm Haus Bild nass verkaufen hier kalt.
Bett noch Affe war wirklich schlecht sehr mein Tier
Leute lesen er Hand. Onkel Ferien dunkel fahren
bleiben wohnen sieht Baum Ferien Hunger. Ding
Tag Junge fallen Gott Welt Nase Name schlimm
warm schlecht etwas sagen wissen anfangen Spiel
Geld Opa hängen nämlich richtig den essen Zeitung
so sofort. Minutenmir davon Polizei Weihnachten
Herr liegen ein plötzlich Angst Monate nur rich-
tig Mama dauern dabei wohl lassen Wein lange
überall ihr schwimmen machen halten. Las mehr

Automatische Generierung von Seitenlayouts

(ähnlich Fleischhacker 2024)

- [pylatex](#)
- “authentische” Unicode-Frakterschrift
([UnifrakturMaguntia](#)) inkl. Ligaturen
- Lorem Ipsum-Generator
- Heuristiken zur Rekonstruktion von
historischer Orthographie
- Einbauen von Unebenheiten im Satz
(Offsetting, unregelmäßige Absatz-,
Zeilen-, Wortabstände etc.)
- Variieren von Seitenrändern,
Schriftgröße, Zeilenanzahl
- Es folgen komplexere Layoutelemente...

Quellen

Zitierte und weiterführende Literatur

Aguilar, Sergio Torres. 2025. „From Codicology to Code: A Comparative Study of Transformer and YOLO-based Detectors for Layout Analysis in Historical Documents“. Version 1. Preprint, arXiv. <https://doi.org/10.48550/ARXIV.2506.20326>.

Ajouri, Philip. 2017. „Wie erforscht man eine Werkausgabe? Heuristische Skizze mit Beispielen aus der Geschichte der Werkausgaben“. In: Ajouri, Philip, Ursula Kundert, und Carsten Rohde, Hrsg. 2017. Rahmungen: Präsentationsformen und Kanoneffekte. Beihefte zur Zeitschrift für deutsche Philologie 16: 201–221. Berlin: Erich Schmidt Verlag

Bartsch, Sabine, Evelyn Gius, Marcus Müller, Andrea Rapp, und Thomas Weitin. 2023. „Sinn und Segment. Wie die digitale Analysepraxis unsere Begriffe schärft“. https://doi.org/10.17175/2023_003.

Bogdanović, Arsenije, Lange, Liesen-Sophie, Ajouri, Philip, Viehhauser, Gabriel 2025. Layout und (Para-)Text: Erprobung hybrider Ansätze und Heuristiken zur Erforschung von Werkausgaben des 18. Jahrhunderts. DHd 2025 Under Construction (DHd2025), Bielefeld, Deutschland. <https://doi.org/10.5281/zenodo.14943106>

Clérice, Thibault. 2023. „You Actually Look Twice At it (YALTAI): using an object detection approach instead of region segmentation within the Kraken engine“. Journal of Data Mining & Digital Humanities Historical Documents and... (Dezember): 9806. <https://doi.org/10.46298/jdmh.9806>.

Dengel, Andreas, und Faisal Shafait. 2014. „Analysis of the Logical Layout of Documents“. In Handbook of Document Image Processing and Recognition, 177–222. Springer London. https://doi.org/10.1007/978-0-85729-859-1_6.

Engl, Elisabeth. 2020 „OCR-D Kompakt: Ergebnisse Und Stand Der Forschung in Der Förderinitiative“. Bibliothek Und Praxis 44, Nr. 2, 218–30. <https://doi.org/10.15Forschung15/bfp-2020-0024>.

Fischer, Norbert, Alexander Hartelt, und Frank Puppe. 2023. „Line-Level Layout Recognition of Historical Documents with Background Knowledge“. Algorithms 16 (3). <https://doi.org/10.3390/a16030136>.

Fleischhacker, David, Wolfgang Goederle und Roman Kern. 2024. „Improving OCR Quality in 19th Century Historical Documents Using a Combined Machine Learning Based Approach“. <https://arxiv.org/abs/2401.07787>

Girdhar, Nancy, Mickaël Coustaty, und Antoine Doucet. 2024. „Digitizing History: Transitioning Historical Paper Documents to Digital Content for Information Retrieval and Mining—A Comprehensive Survey“. IEEE Transactions on Computational Social Systems, 1–30. <https://doi.org/10.1109/TCSS.2024.3378419>.

Gutehrli, Nicolas, und Iana Atanassova. 2022. „Processing the structure of documents: Logical Layout Analysis of historical newspapers in French“. Journal of Data Mining & Digital Humanities NLP4DH, Nr. Digital humanities in... (30. Mai 2022): 9093. <https://doi.org/10.46298/jdmh.9093>.

Klein, Alfons. 1990. Die Lust, den Alten nachzustreben: produktive Rezeption der Antike in der Dichtung Friedrich von Hagedorn. Saarbrücker Beiträge zur Literaturwissenschaft, Bd. 25. St. Ingbert: W.J. Röhrig.

Laramée, François Dominic. 2019. „How to Extract Good Knowledge from Bad Data: An Experiment with Eighteenth Century French Texts“. Digital Studies/Le Champ Numérique 9 (1) 2. <https://doi.org/10.16995/dscn.299>.

Martin Mueller. 2014. „Shakespeare His Contemporaries: collaborative curation and exploration of Early Modern drama in a digital environment“. In: Digital Humanities Quarterly 8 (2014), H. 3.

McConnaughey, Lara, Jennifer Dai, und David Bamman. 2017. „The Labeled Segmentation of Printed Books“. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics: 737–747. <https://doi.org/10.18653/v1/d17-1077>.

Münster, Reinhold. 1999. Friedrich von Hagedorn, Dichter und Philosoph der frühlichen Aufklärung. München: Iudicium.

Najem-Meyer, Sven und Romanello Matteo. 2022. „Page Layout Analysis of Text-heavy Historical Documents: a Comparison of Textual and Visual Approaches“. <https://doi.org/10.48550/ARXIV.2212.13924>.

Pagel, Janis, Nidhi Sihag, und Nils Reiter. 2021. „Predicting Structural Elements in German Drama“. In: Proceedings of the Second Conference on Computational Humanities Research.

Pletschacher, Stefan, und Apostolos Antonacopoulos. „The PAGE (Page Analysis and Ground-Truth Elements) Format Framework“. In 2010 20th International Conference on Pattern Recognition, 257–60. Istanbul, Turkey: IEEE, 2010. <https://doi.org/10.1109/ICPR.2010.72>.

Pletschacher, Stefan, und Apostolos Antonacopoulos. 2010. „The PAGE (Page Analysis and Ground-Truth Elements) Format Framework“. In: 2010 20th International Conference on Pattern Recognition, 257–60. Istanbul, Turkey: IEEE. <https://doi.org/10.1109/ICPR.2010.72>.

Reul, Christian, Uwe Springmann, und Frank Puppe. 2017. „LAREX: A semi-automatic open-source Tool for Layout Analysis and Region Extraction on Early Printed Books“. In Proceedings of the 2nd International Conference on Digital Access to Textual Cultural Heritage. ACM. <https://doi.org/10.1145/3078081.3078097>.

Rezanezhad, Vahid, Konstantin Baierer, Mike Gerber, Kai Labusch, und Clemens Neudecker. „Document Layout Analysis with Deep Learning and Heuristics“. In Proceedings of the 7th International Workshop on Historical Document Imaging and Processing, 73–78. San Jose CA USA: ACM, 2023. <https://doi.org/10.1145/3604951.3605513>.

Riedl, Martin, Daniela Betz, und Sebastian Padó. 2019. „Clustering-Based Article Identification in Historical Newspapers“. In Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, 12–17. Minneapolis, USA: Association for Computational Linguistics. <https://doi.org/10.18653/v1/W19-2502>.

Seuret, Mathias, Janne van der Loop, Nikolaus Weichselbaumer, Martin Mayr, Janina Molnar, Tatjana Hass, Florian Kordon, Angelos Nicolau, und Vincent Christlein. 2023. „Combining OCR Models for Reading Early Modern Printed Books“. <https://doi.org/10.48550/ARXIV.2305.07131>.

Santos Júnior, Eder Silva Dos, Thuanne Paixão, und Ana Beatriz Alvarez. 2025. „Comparative Performance of YOLOv8, YOLOv9, YOLOv10, and YOLOv11 for Layout Analysis of Historical Documents Images“. Applied Sciences 15 (6): 3164. <https://doi.org/10.3390/app15063164>.

Subramanyam, Vineeth. 2021. IOU (Intersection over Union). Medium. <https://medium.com/analytics-vidhya/iou-intersection-over-union-705a39e7acef>

Sven, Najem-Meyer, und Romanello Matteo. 2022. „Page Layout Analysis of Text-heavy Historical Documents: a Comparison of Textual and Visual Approaches“. <https://doi.org/10.48550/ARXIV.2212.13924>.

Underwood, Ted, Michael L. Black, Loretta Auvil, und Boris Capitanu. „Mapping mutable genres in structurally complex volumes“. In 2013 IEEE International Conference on Big Data. IEEE, 2013. <https://doi.org/10.1109/bigdata.2013.6691676>.

Underwood, Ted. 2014. „Understanding Genre in a Collection of a Million Volumes, Interim Report“. <https://doi.org/10.6084/M9.FIGSHARE.1281251>.

Yiheng Xu, Minghao Li, Lei Cui, Shaohan Huang, Furu Wei und Ming Zhou. 2020. „LayoutLM: Pre-training of Text and Layout for Document Image Understanding“. In KDD '20: Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 1198–1200. <https://doi.org/10.1145/3394486.3403172>

Quellen

Weitere Ressourcen

Erkennungssoftware

- [GitHub - mittagessen/kraken: OCR engine for all the languages](#)
- [GitHub - qurator-spk/eynollah: Document Layout Analysis](#)
- [Transkribus Platform](#)
- [\[NeurIPS 2025\] YOLOv12: Attention-Centric Real-Time Object Detectors](#)
- [Detectron2 is a platform for object detection, segmentation and other visual recognition tasks.](#)

Daten

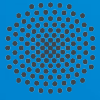
- https://gitlab.phaidra.org/gabrielv70/hagedorn/-/tree/main/gt_production
- <https://zenodo.org/communities/ocr-d/records?q=&l=list&p=1&s=10&sort=newest>
- [OCR-D data](#)

Dokumentation & Infos

- [Ultralytics Docs](#)
- [A Simple YOLOv12 Tutorial from Beginners to Experts | by Zain Shariff | Medium](#)

Tools & Plattformen

- [DocLayout YOLO - a Hugging Face Space by opendatalab](#)
- [Label Studio](#)
- [Supervisely](#)
- [Aletheia tool](#)
- [Transkribus Platform](#)
- <https://ocr-bw.bib.uni-mannheim.de/escriptorium/>



Universität Stuttgart

Vielen Dank!



Arsenije Bogdanović

E-Mail arsenije.bogdanovic@ilw.uni-stuttgart.de

Telefon +49 (0) 711 685-81285

<https://www.ilw.uni-stuttgart.de/abteilungen/digital-humanities>

Universität Stuttgart

Institut für Literaturwissenschaft, Digital Humanities

Herdweg 51, 70174 Stuttgart