Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

OCR4all –
Eine (semi)automatische Open Source Software für die OCR historischer Drucke

Christian Reul

Zentrum für Philologie und Digitalität „Kallimachos“ (ZPD)
Universität Würzburg

05.05.2021

Gliederung

1. **Einleitung**
2. Submodule
3. Workflow
4. Live Demo
5. Evaluation
6. Diskussion und Ausblick

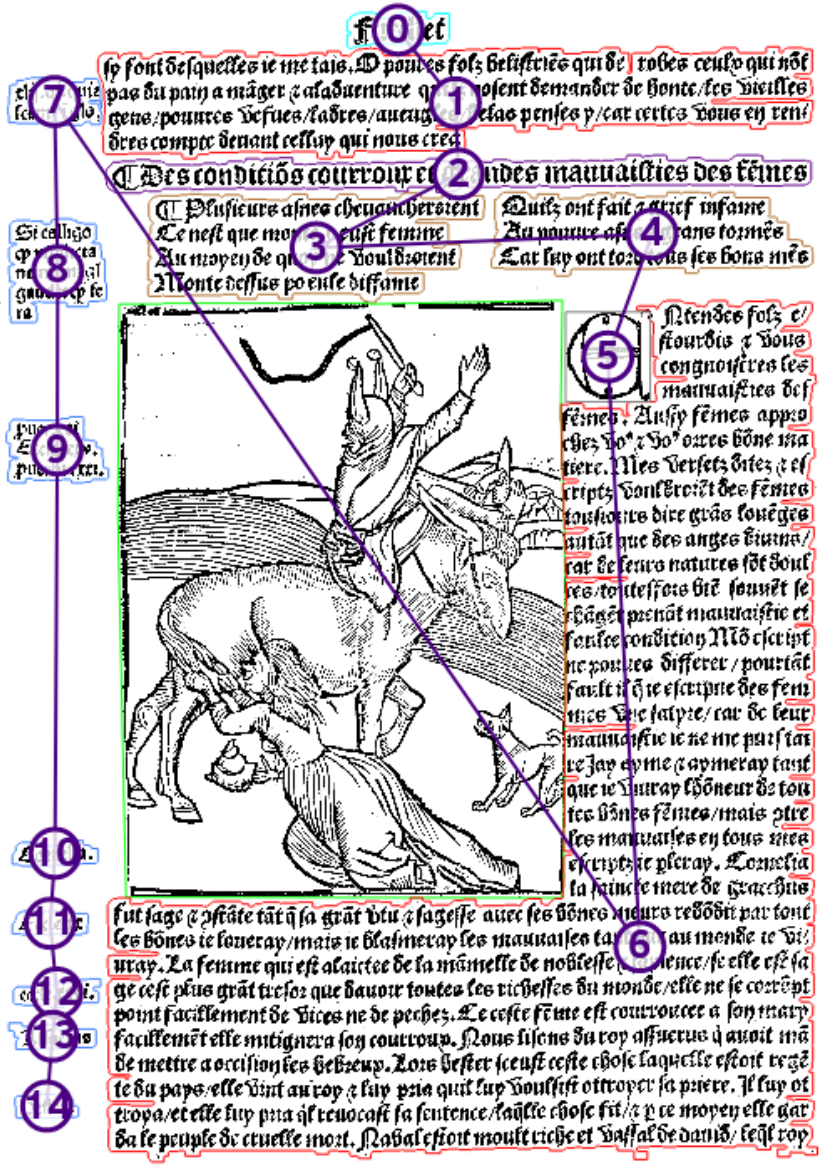
OCR4all – Motivation und Überblick

- Bestehende Open Source OCR Software ist mächtig, kann nicht-technische Nutzer jedoch schnell überfordern:
  - Installation nicht unbedingt trivial
  - Keine grafische Benutzeroberfläche, sondern ungewohnte Kommandozeile
  - …
- Die Idee hinter OCR4all:
  - Verständlich und anwendbar auch für nicht-technische Nutzer
  - Nutzung von Docker/VirtualBox → Einfache Installation, Unabhängigkeit vom Betriebssystem
  - Basiert auf unterschiedlichen Open Source OCR Tools

OCR4all – Historie und Aktueller Stand

- Ursprünglich entwickelt für die OCR sehr (sehr!) alter Drucke (Projekt Narragonien digital):
  - Sehr komplizierte Layouttypisierung
  - Training werkspezifischer Modelle unerlässlich aufgrund der höchst varianten Typographie
  - Anspruch häufig: 100% Erkennungsgenauigkeit für Layout und OCR

  → Nutzer nehmen dafür einen gewissen manuellen Korrekturaufwand in Kauf
- Mittlerweile erfolgreich auf großer Bandbreite von Drucken (15. bis 21. Jh.) eingesetzt
- Hauptziel: Erhöhung des Automatisierungsgrads und der Robustheit
- Work in progress!



www.ocr4all.de

Gliederung

1. Einleitung
2. **Submodule**
3. Workflow
4. Live Demo
5. Evaluation
6. Diskussion und Ausblick

Aktuell integriert

- OCRopus
  - Vorverarbeitung
  - Automatische (Zeilen)Segmentierung
- Calamari
  - Texterkennung
  - Modelltraining
  - Evaluation
- LAREX
  - Ursprünglich zur interaktiven Regionensegmentierung inklusive semantischer Auszeichnung entwickelt
  - Mittlerweile umfangreiches Korrekturtool, u. a. für
    - Regionen- und Zeilenkoordinaten
    - Semantische Typisierung und Lesereihenfolge
    - Text (Ground Truth Erstellung für Training!)

Zwischenfazit und -ausblick

- Bislang begrenzte Auswahl
  - Ursprünglich für konkreten Anwendungsfall (Narragonien digital) entwickelt
  - Damaliges Ziel: erstmal funktionierenden Workflow schaffen
  - Fokus auf ausgewählte „beste“ Lösungen
- Anbieten weiterer Lösungen unbedingt intendiert
  - Problemlose Anbindung, solange klar definierte Schnittstellen eingehalten werden
  - Workflow und Lösungen für einzelne Schritte frei kombinierbar
  - Baukastenprinzip!
- Klingt nach OCR-D?!

  Später mehr…

„Nimm die Binarisierung von OCRopus, die Segmentierung von Tesseract und die Texterkennung von Calamari“

(Konstantin Baierer, vor 20min)

Gliederung

1. Einleitung
2. Submodule
3. **Workflow**
4. Live Demo
5. Evaluation
6. Diskussion und Ausblick

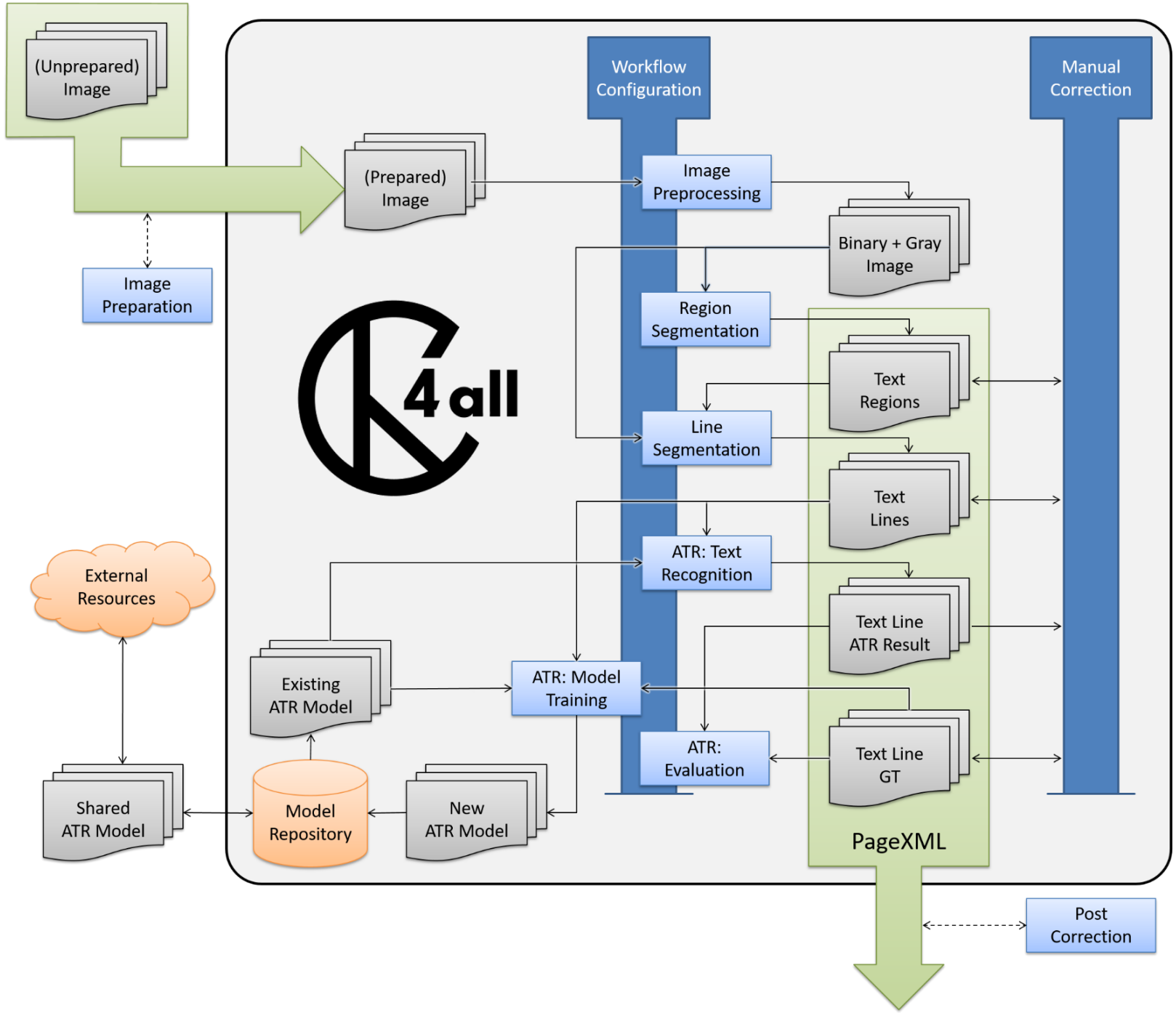

(Unprepared) Image
Image Preparation
(Prepared) Image
Workflow Configuration
Manual Correction
Image Preprocessing
Binary + Gray Image
Region Segmentation
Text Regions
Line Segmentation
Text Lines
ATR: Text Recognition
Text Line ATR Result
ATR: Model Training
ATR: Evaluation
Text Line GT
PageXML
External Resources
Existing ATR Model
Shared ATR Model
Model Repository
New ATR Model
Post Correction

Image Preparation

- **Input:** unverarbeitete Bilder
  Output: vorbereitete Bilder
- Notwendige Vorbereitungen variieren von Werk zu Werk:
  - Teilung von Doppelseiten
  - Rotation
  - Entfernung von störenden Bildrändern
  - …
- Bisher nicht in OCR4all integriert
  → Open Source Tool [ScanTailor]

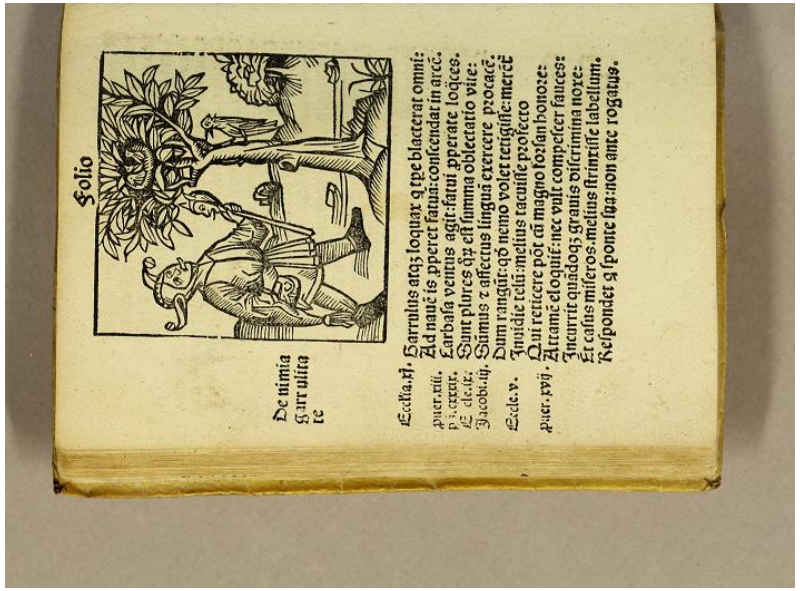
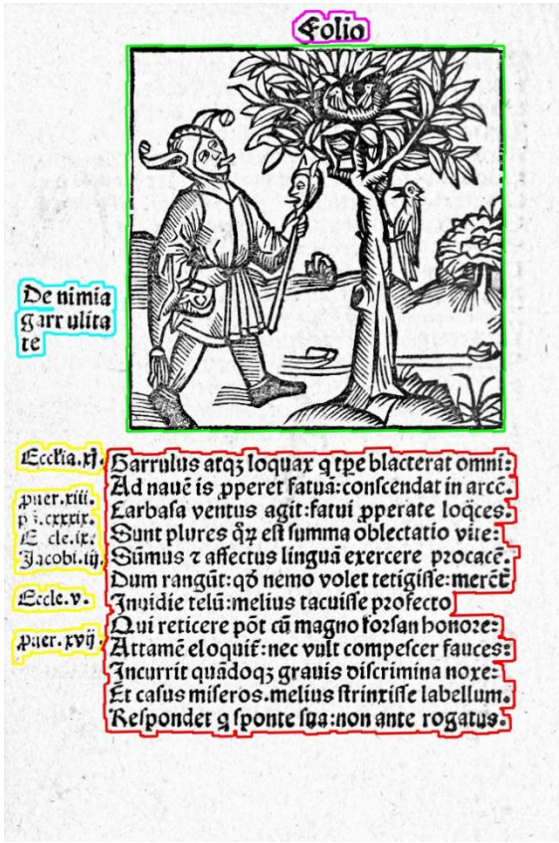


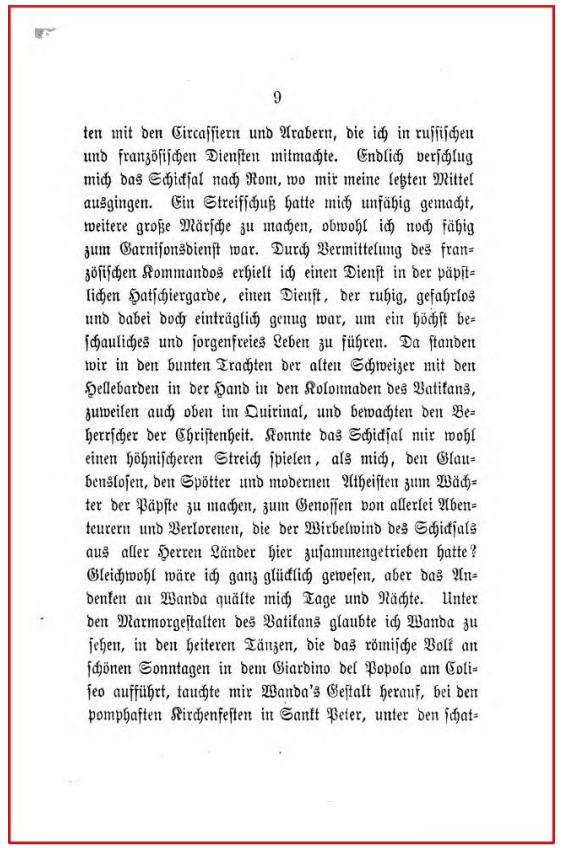
Image Preprocessing

- **Input:** Originalbilder (Farb- oder Graustufen- oder Binärbilder)
  Output: geradegestellte Binärbilder
- Zwei Teilschritte, die die folgenden Arbeitsschritte erleichtern:
  - Binarisierung
  - Geradestellen
- Derzeit umgesetzt durch *ocropus-nlbin*

Region Segmentation

- **Input:** vorverarbeitete Bilder
  Output: Informationen über vorhandene Layoutelemente und deren Lesereihenfolge
- Tools/Methoden:
  - LAREX
    - Exakte Segmentierung inklusive semantischer Auszeichnung
    - Semiautomatisch, Intuitiv, adaptierbar und nachvollziehbar
  - Dummy Segmentation
    - Ganze Seite als ein Textsegment, Rest erledigt Zeilensegmentierung
    - Vollautomatisch und sehr schnell
    - Oft völlig ausreichend für moderate Layouts (z. B. typische Fraktur Romane aus dem 19. Jh.)





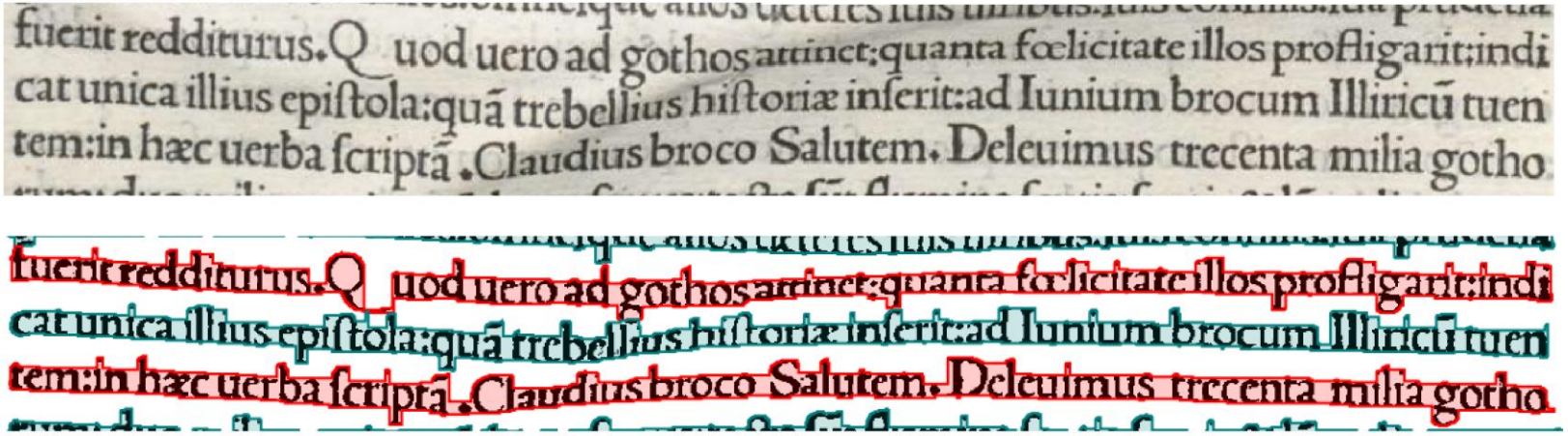
9

ten mit den Circaſſiern und Arabern, die ich in ruſſiſchen und franzöſiſchen Dienſten mitmachte. Endlich verſchlug mich das Schickſal nach Rom, wo mir meine letzten Mittel ausgingen. Ein Streifſchuß hatte mich unfähig gemacht, weitere große Märſche zu machen, obwohl ich noch fähig zum Garniſonsdienſt war. Durch Vermittelung des franzöſiſchen Kommandos erhielt ich einen Dienſt in der päpſtlichen Hatſchiergarde, einen Dienſt, der ruhig, gefahrlos und dabei doch einträglich genug war, um ein höchſt beſchauliches und ſorgenfreies Leben zu führen. Da ſtanden wir in den bunten Trachten der alten Schweizer mit den Hellebarden in der Hand in den Kolonnaden des Vatikans, zuweilen auch oben im Quirinal, und bewachten den Beherrſcher der Chriſtenheit. Konnte das Schickſal mir wohl einen höhniſcheren Streich ſpielen, als mich, den Glaubensloſen, den Spötter und modernen Atheiſten zum Wächter der Päpſte zu machen, zum Genoſſen von allerlei Abenteurern und Verlorenen, die der Wirbelwind des Schickſals aus aller Herren Länder hier zuſammengetrieben hatte? Gleichwohl wäre ich ganz glücklich geweſen, aber das Andenken an Wanda quälte mich Tage und Nächte. Unter den Marmorgeſtalten des Vatikans glaubte ich Wanda zu ſehen, in den heiteren Tänzen, die das römiſche Volk an ſchönen Sonntagen in dem Giardino del Popolo am Coliſeo aufführt, tauchte mir Wanda's Geſtalt herauf, bei den pomphaften Kirchenfeſten in Sankt Peter, unter den ſchat-

Line Segmentation

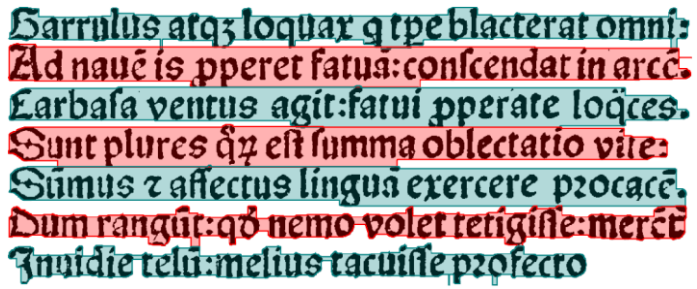
- **Input:** vorverarbeitete Bilder und Segmentierungsinformationen
  Output: Zeilenpolygone
- Wichtiger Vorbereitungsschritt der eigentlichen Texterkennung

- Vorab Deskewing einzelner Regionen
- Zeilensegmentierung auch bei problematischer Ausgangslage möglich
- Derzeit umgesetzt durch angepasstes *ocropus-gpageseg*

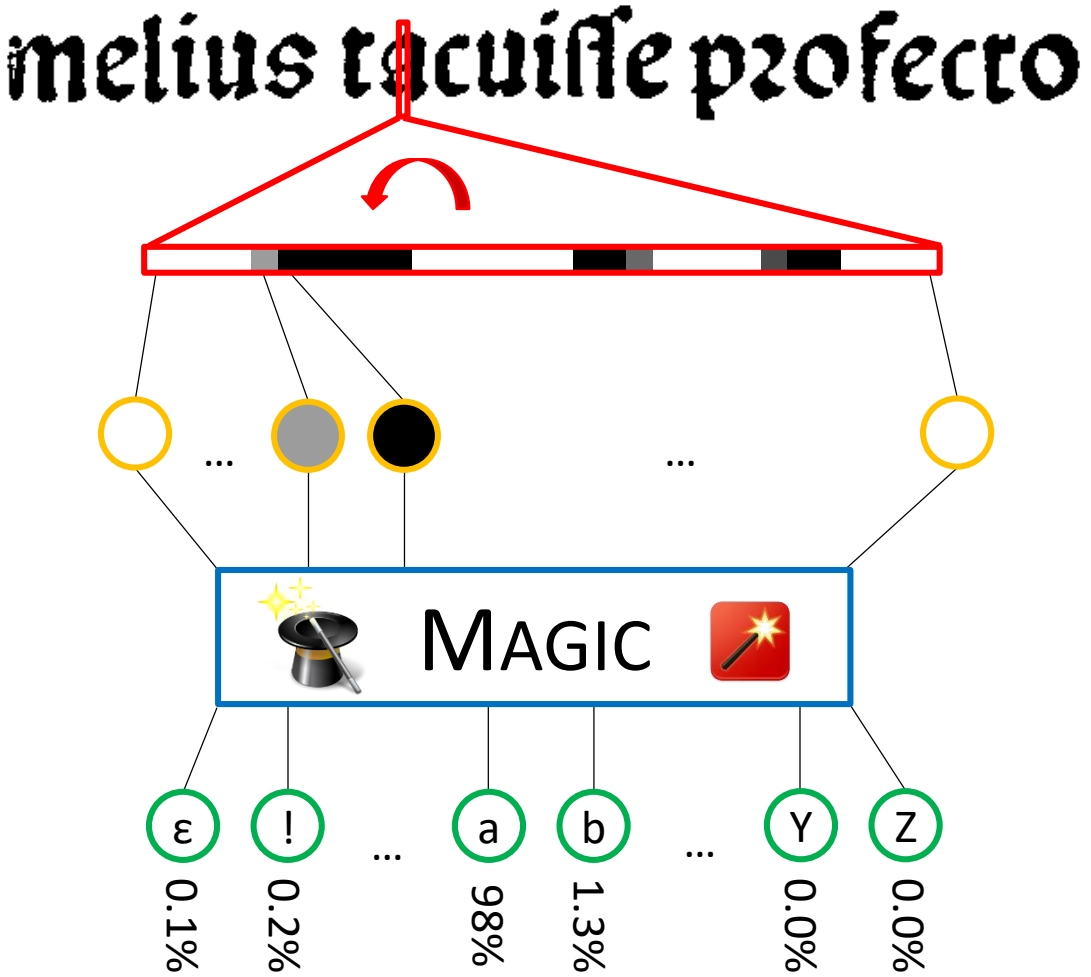


Character Recognition

- **Input:** Textzeilenbilder und OCR-Modelle
  Output: erkannte Textzeilen
- Derzeit umgesetzt durch *calamari-predict*

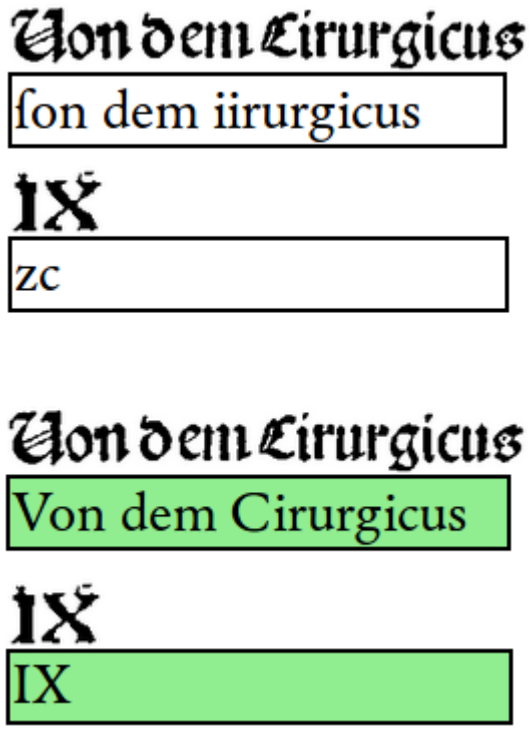


Earrulus atqȝ loquax q Epe blacterat omni:
Ad nauē is pperet fatuā:conſcenda in arcē
Sunt plures qu eſt ſurmma oblectatio vite:
Sūmus ꝛ affectus linguſta exercere pꝛocacē.
Dum rangūt:qꝺ nemo volet tetigiſſe:merct
Jnuidie telū:melius tacuiſſe pꝛofecto

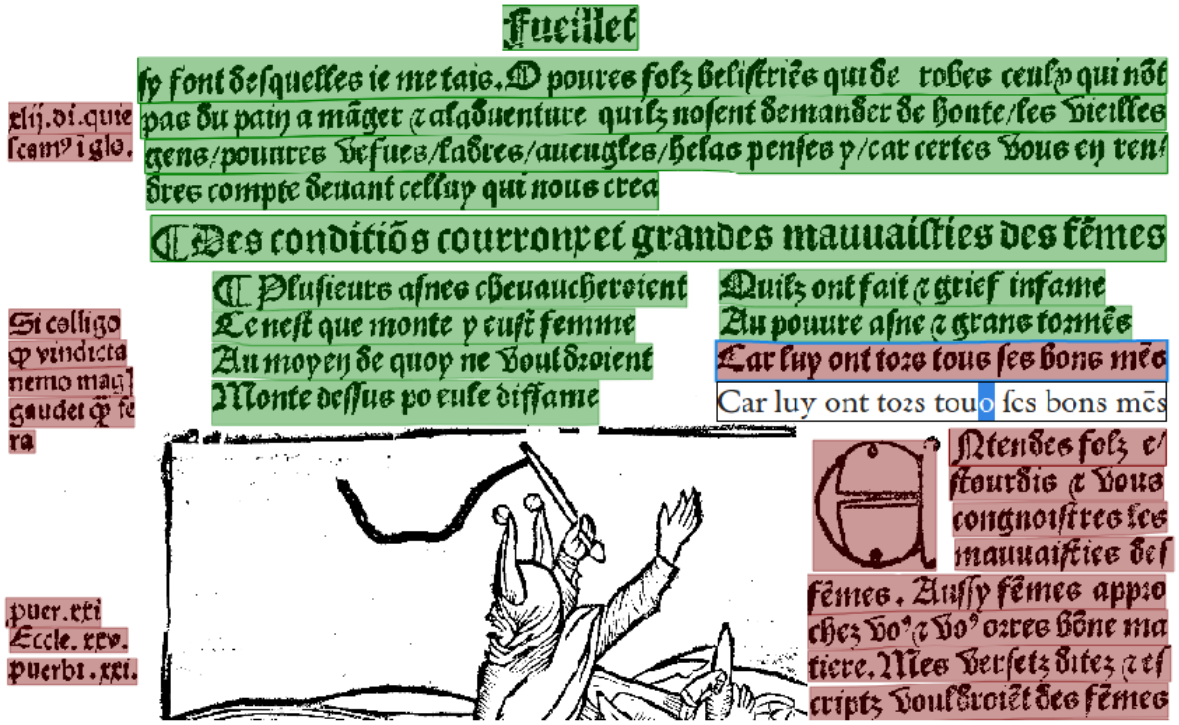


Post Correction – Ground Truth Production

- **Input:** Textzeilenbilder und Erkennungsergebnisse
  Output: korrigierter Text (= Ground Truth, GT)
- GT wird für das Training von OCR-Modellen benötigt
- Anpassbares Virtuelles Keyboard ermöglicht die Verwendung spezifischer Sonderzeichen



Post Correction – LAREX



Quilz ont fait ⁊ grief infame

Ouilz ont fait ⁊ grief infame

Au pouure asne ⁊ grans tozmẽs

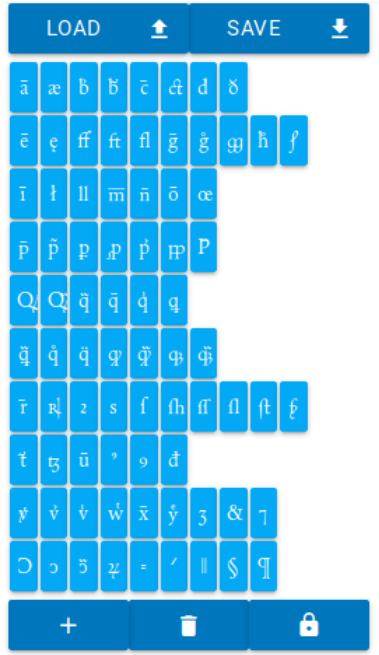
Au pouure asne ⁊ grans tozmẽs

Car luy ont tozs tous ses bons mẽs

Car luy ont tozs touo ſes bons mẽs

Ntendes folz e/

Ntendes folz e=



Evaluation

- **Input:** OCR-Ergebnisse und Ground Truth für bestimmte Textzeilen
  Output: Zeichenfehlerrate (CER) und -statistiken
- Zeigt die Art und Häufigkeit der auftretenden Fehler an
- Auf Grundlage der Evaluation lässt sich erkennen, ob Fehler (noch) systematisch sind oder nicht
- Derzeit umgesetzt durch *calamari-eval*

GT	PRED	COUNT	PERCENT
{ꝛ}	{r}	33	17.65%
{}	{ }	11	5.88%
{ī}	{i}	11	5.88%
{ﬄ}	{f}	7	3.74%
{ﬃ}	{ff}	5	5.35%
{t}	{i}	3	1.60%
{q}	{}	2	1.07%
{◌̊}	{°}	2	1.07%
{ }	{}	2	1.07%
{x}	{}	2	1.07%

Training

Umfangreiche Unterstützung …

- … der zahlreichen Möglichkeiten des Calamari Trainingsprozesses
  - Ensemble-Training
  - Pretraining/Finetuning
  - Datenaugmentierung
  - …
- … des iterativen Trainingsansatzes
  - Ständiges Durchlaufen der Schritte Erkennung, Korrektur und Training zur Effizienzsteigerung
  - Modellverwaltung

Für Details siehe Veranstaltung zum Thema „Training“ nächste Woche.

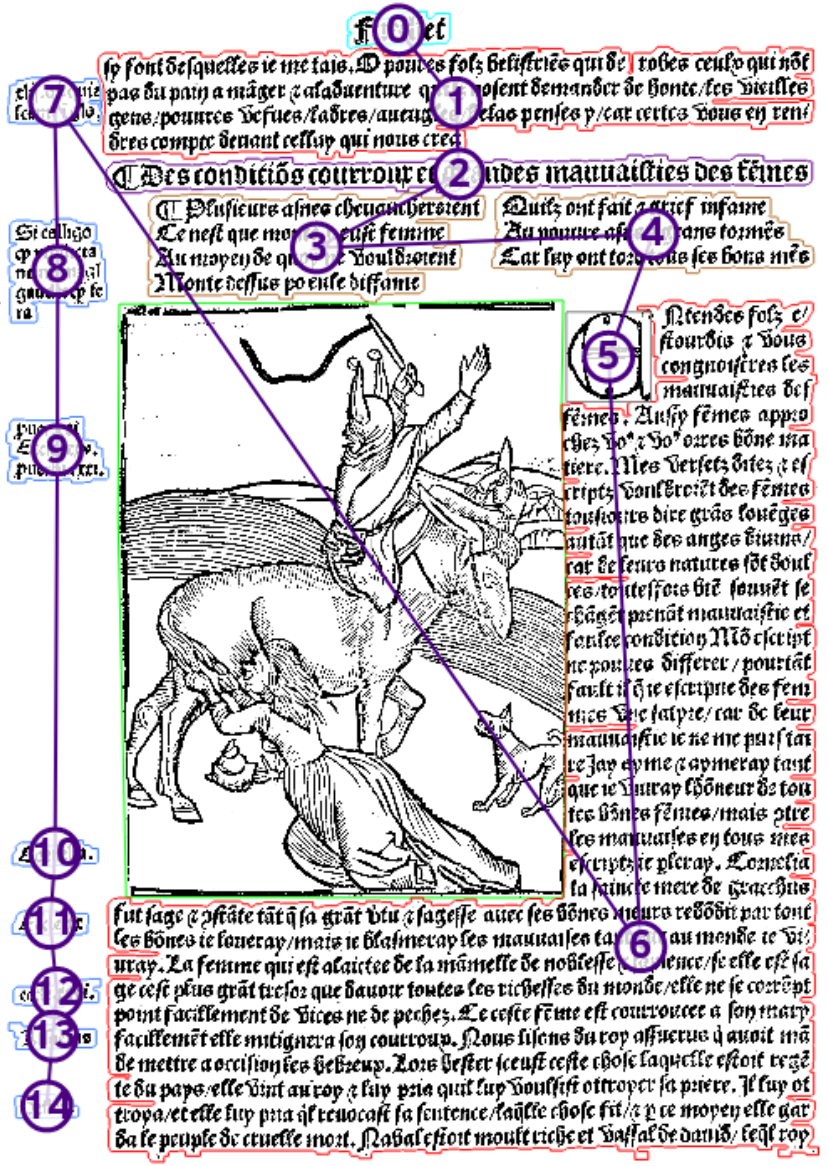
Live Demo!

Gliederung

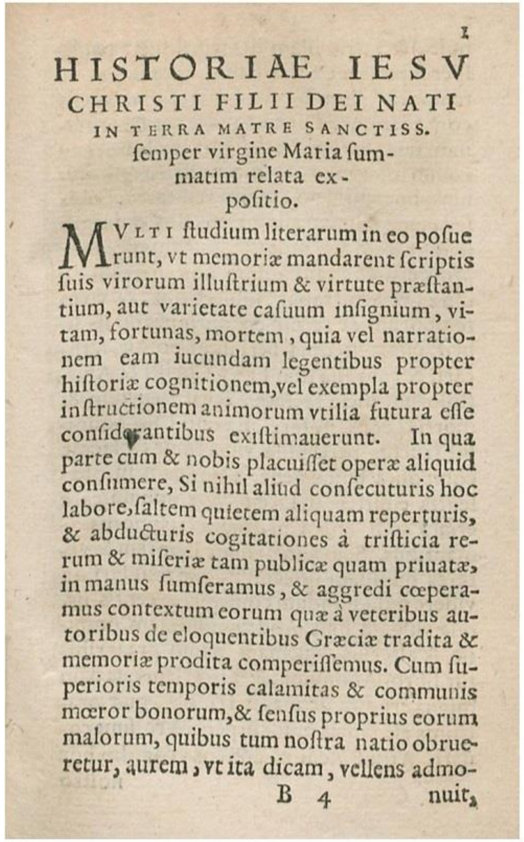
1. Einleitung
2. Submodule
3. Workflow
4. Live Demo
5. **Evaluation**
6. Diskussion und Ausblick

Evaluation – Frühdrucke bis 1600

- Material
  - 5 Ausgaben des „Narrenschiffs“
  - 17 Werke des Universalgelehrten Joachim Camerarius
  - 3 gemischte, früh-neuzeitliche Drucke (Praktikum)
- Anforderungen / Ziele:
  - Fehlerfreie Segmentierung und Reading Order sowie präzise semantische Auszeichnung
  - Final: zitierfähiger Volltext (0% Fehler)
  - Vorerst: 1% CER oder besser
- Bearbeitung in zwei Gruppen:
  - Unerfahrene, nicht-technische Nutzer
  - Erfahrene Nutzer



Frühdrucke bis 1600 – Beispiele I

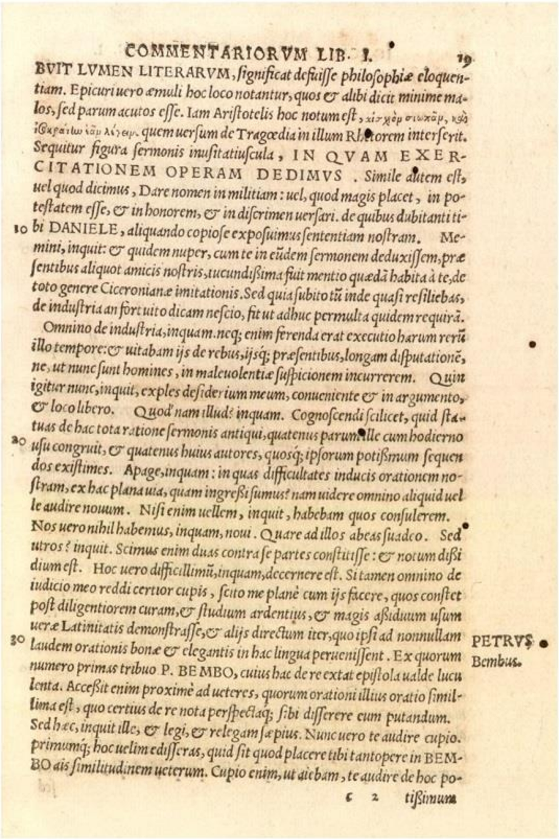


1

HISTORIAE IESV
CHRISTI FILII DEI NATI
IN TERRA MATRE SANCTISS.
ſemper virgine Maria ſum-
matim relata ex-
poſitio.

MVLTI ſtudium literarum in eo poſuerunt, vt memoriæ mandarent ſcriptis ſuis virorum illuſtrium & virtute præſtantium, aut varietate caſuum inſignium, vitam, fortunas, mortem, quia vel narrationem eam iucundam legentibus propter hiſtoriæ cognitionem, vel exempla propter inſtructionem animorum vtilia futura eſſe conſiderantibus exiſtimauerunt. In qua parte cum & nobis placuiſſet operæ aliquid conſumere, Si nihil aliud conſecuturis hoc labore, ſaltem quietem aliquam reperturis, & abducturis cogitationes à triſticia rerum & miſeriæ tam publicæ quam priuatæ, in manus ſumſeramus, & aggredi cœperamus contextum eorum quæ à veteribus autoribus de eloquentibus Græciæ tradita & memoriæ prodita comperiſſemus. Cum ſuperioris temporis calamitas & communis mœror bonorum, & ſenſus proprius eorum malorum, quibus tum noſtra natio obrueretur, aurem, vt ita dicam, vellens admo-

B 4 nuit,



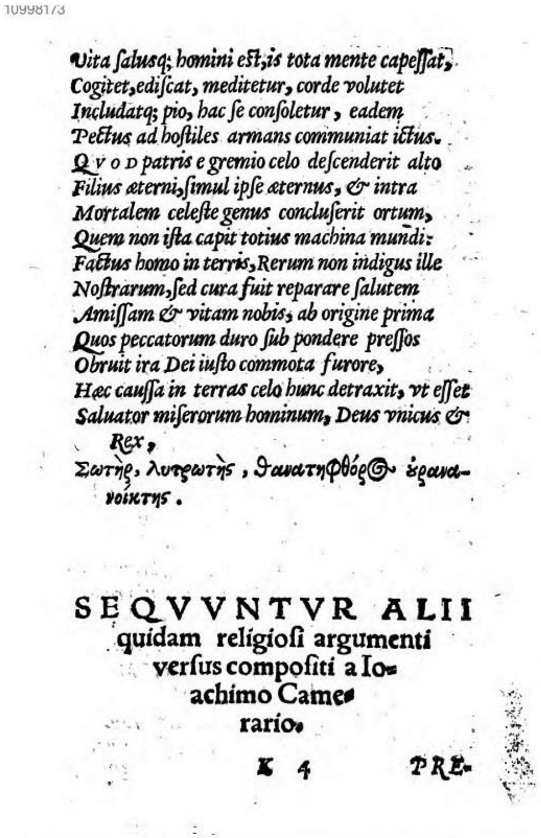
COMMENTARIORVM LIB. I. 19

BVIT LVMEN LITERARVM, ſignificat defuiſſe philoſophiæ eloquentiam. Epicuri uero æmuli hoc loco notantur, quos & alibi dicit minime malos, ſed parum acutos eſſe. Iam Ariſtotelis hoc notum eſt, αἰσχρὸν σιωπᾶν, καὶ Ἰσοκράτην ἐᾶν λέγειν. quem uerſum de Tragœdia in illum Rhetorem interſerit. Sequitur figura ſermonis inuſitatiuſcula, IN QVAM EXERCITATIONEM OPERAM DEDIMVS. Simile autem eſt, uel quod dicimus, Dare nomen in militiam: uel, quod magis placet, in poteſtatem eſſe, & in honorem, & in diſcrimen uerſari. de quibus dubitanti tibi DANIELE, aliquando copioſe expoſuimus ſententiam noſtram. Memini, inquit: & quidem nuper, cum te in eũdem ſermonem deduxiſſem, præſentibus aliquot amicis noſtris, iucundiſsima fuit mentio quædã habita à te, de toto genere Ciceronianæ imitationis. Sed quia ſubito tũ inde quaſi reſiliebas, de induſtria an fortuito dicam neſcio, fit ut adhuc permulta quidem requirã.

Omnino de induſtria, inquam. neq; enim ferenda erat executio harum rerũ illo tempore: & uitabam ijs de rebus, ijsq; præſentibus, longam diſputationẽ, ne, ut nunc ſunt homines, in maleuolentiæ ſuſpicionem incurrerem. Quin igitur nunc, inquit, exples deſiderium meum, conueniente & in argumento, & loco libero. Quod nam illud? inquam. Cognoſcendi ſcilicet, quid ſtatuas de hac tota ratione ſermonis antiqui, quatenus parum ille cum hodierno uſu congruit, & quatenus huius autores, quosq; ipſorum potiſsimum ſequendos exiſtimes. Apage, inquam: in quas difficultates inducis orationem noſtram, ex hac plana uia, quam ingreſsi ſumus? nam uidere omnino aliquid uelle audire nouum. Niſi enim uellem, inquit, habebam quos conſulerem. Nos uero nihil habemus, inquam, noui. Quare ad illos abeas ſuadeo. Sed utros? inquit. Scimus enim duas contra ſe partes conſtitiſſe: & notum diſsidium eſt. Hoc uero difficillimũ, inquam, decernere eſt. Si tamen omnino de iudicio meo reddi certior cupis, ſcito me planè cum ijs facere, quos conſtet poſt diligentiorem curam, & ſtudium ardentius, & magis aſsiduum uſum ueræ Latinitatis demonſtraſſe, & alijs directum iter, quo ipſi ad nonnullam laudem orationis bonæ & elegantis in hac lingua perueniſſent. Ex quorum numero primas tribuo P. BEMBO, cuius hac de re extat epiſtola ualde luculenta. Acceſsit enim proximè ad ueteres, quorum orationi illius oratio ſimillima eſt, quo certius de re nota perſpectaq; ſibi differere eum putandum. Sed hæc, inquit ille, & legi, & relegam ſæpius. Nunc uero te audire cupio. primumq; hoc uelim ediſſeras, quid ſit quod placere tibi tantopere in BEMBO ais ſimilitudinem ueterum. Cupio enim, ut aiebam, te audire de hoc po-

PETRVS Bembus.

c 2 tiſsimum

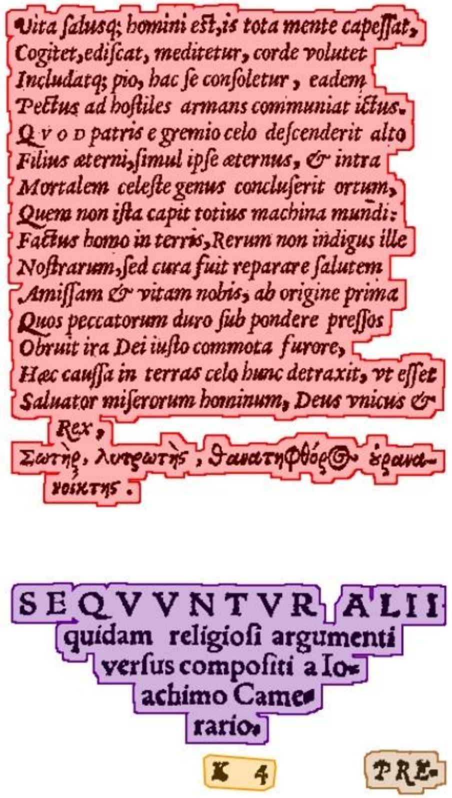


10998173

Vita ſalusq; homini eſt, is tota mente capeſſat,
Cogitet, ediſcat, meditetur, corde volutet
Includatq; pio, hac ſe conſoletur, eadem
Pectus ad hoſtiles armans communiat ictus.
QVOD patris e gremio celo deſcenderit alto
Filius æterni, ſimul ipſe æternus, & intra
Mortalem celeſte genus concluſerit ortum,
Quem non iſta capit totius machina mundi:
Factus homo in terris, Rerum non indigus ille
Noſtrarum, ſed cura fuit reparare ſalutem
Amiſſam & vitam nobis, ab origine prima
Quos peccatorum duro ſub pondere preſſos
Obruit ira Dei iuſto commota furore,
Hæc cauſſa in terras celo hunc detraxit, vt eſſet
Saluator miſerorum hominum, Deus vnicus &
Rex,
Σωτὴρ, λυτρωτὴς, θανατηφθόρος ἐρανα-
νοίκτης.

SEQVVNTVR ALII
quidam religioſi argumenti
verſus compoſiti a Ioachimo Camerario.

K 4 PRE-

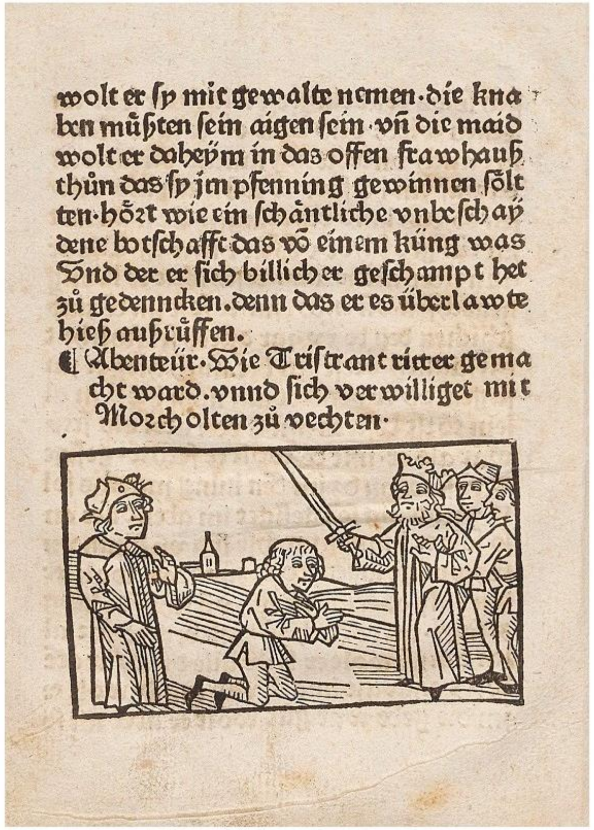


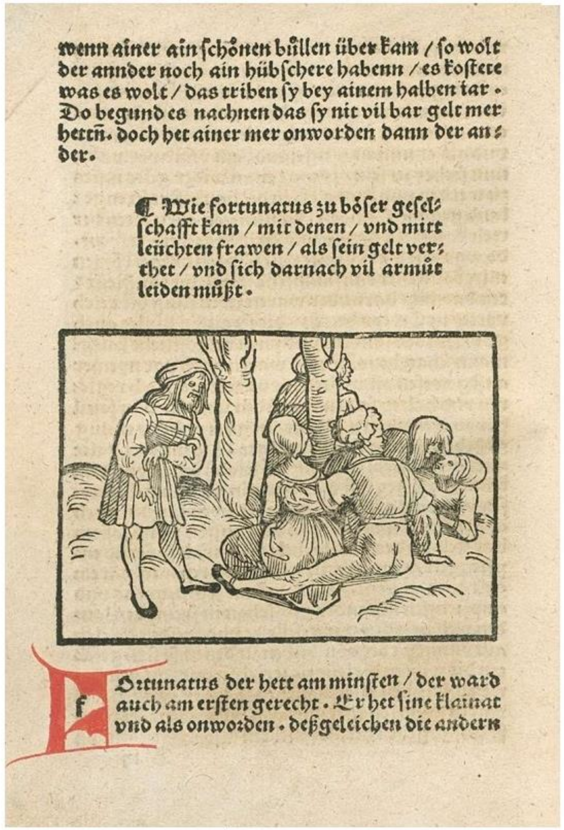
Vita ſalusq; homini eſt, is tota mente capeſſat,
Cogitet, ediſcat, meditetur, corde volutet
Includatq; pio, hac ſe conſoletur, eadem
Pectus ad hoſtiles armans communiat ictus.
QVOD patris e gremio celo deſcenderit alto
Filius æterni, ſimul ipſe æternus, & intra
Mortalem celeſte genus concluſerit ortum,
Quem non iſta capit totius machina mundi:
Factus homo in terris, Rerum non indigus ille
Noſtrarum, ſed cura fuit reparare ſalutem
Amiſſam & vitam nobis, ab origine prima
Quos peccatorum duro ſub pondere preſſos
Obruit ira Dei iuſto commota furore,
Hæc cauſſa in terras celo hunc detraxit, vt eſſet
Saluator miſerorum hominum, Deus vnicus &
Rex,
Σωτὴρ, λυτρωτὴς, θανατηφθόρος ἐρανα-
νοίκτης.

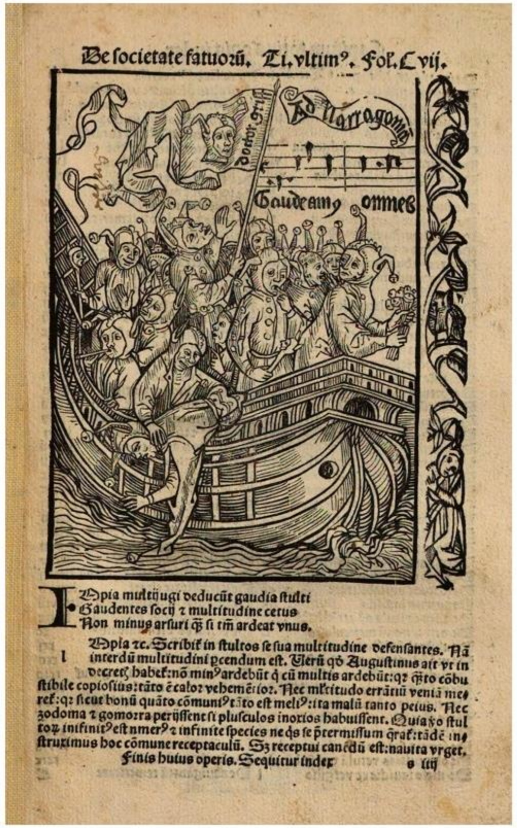
SEQVVNTVR ALII
quidam religioſi argumenti
verſus compoſiti a Ioachimo Camerario.

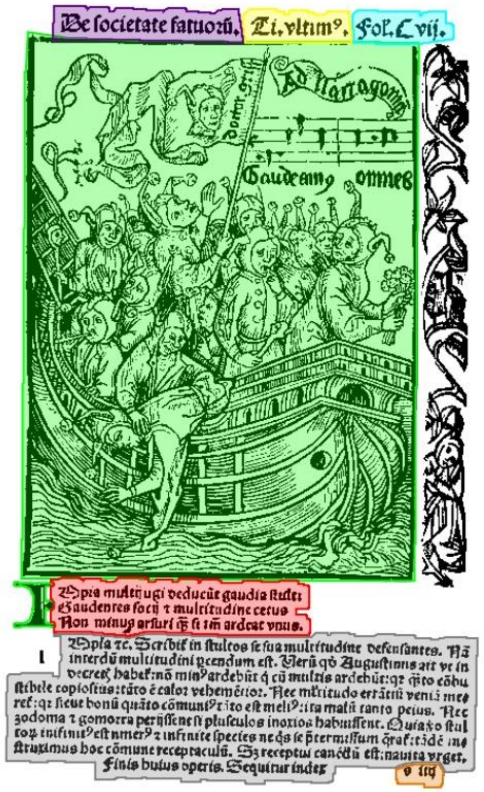
K 4 PRE-

Frühdrucke bis 1600 – Beispiele II

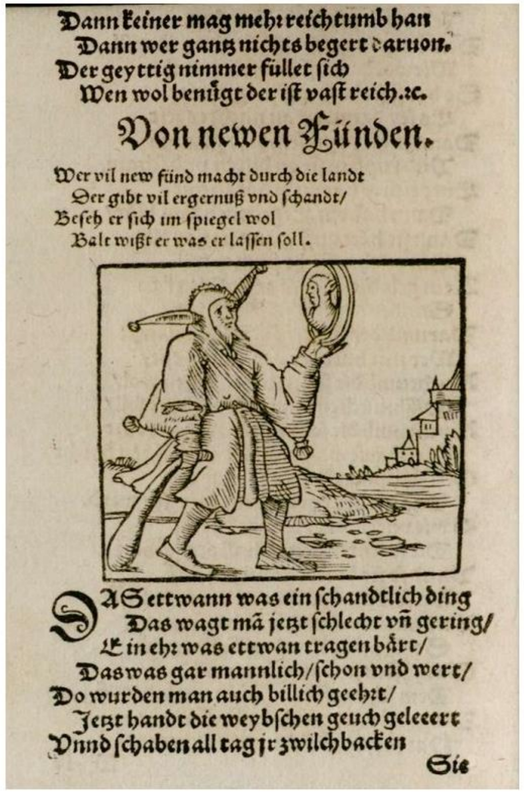






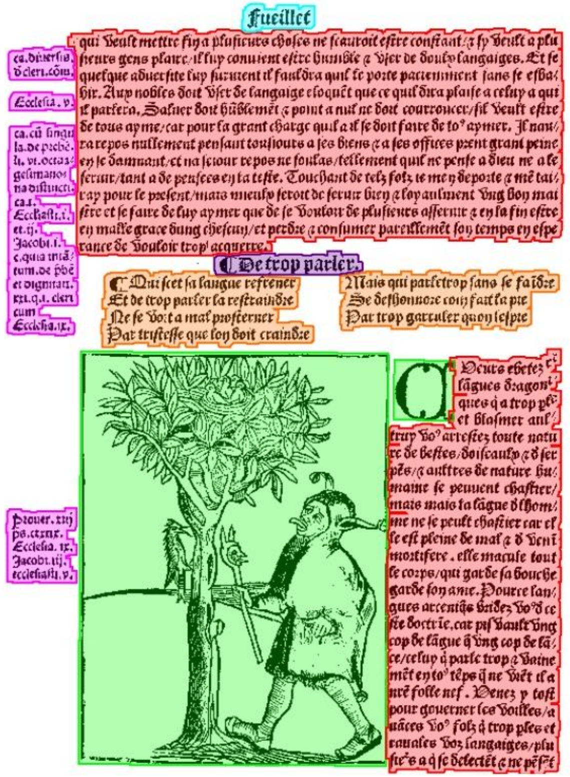


Frühdrucke bis 1600 – Beispiele III


Dann keiner mag mehr reichtumb han
Dann wer gantz nichts begert daruon.
Der geytig nimmer füllet sich
Wen wol benügt der ist vast reich.rc.
Von newen Fünden.
Wer vil new fünd macht durch die landt
Der gibt vil ergernuß vnd schandt/
Beseh er sich im spiegel wol
Balt wißt er was er lassen soll.
DAS ettwann was ein schandtlich ding
Das wagt mã jetzt schlecht vñ gering/
Ein ehr was ettwan tragen bärt/
Das was gar mannlich/schon vnd wert/
Do wurden man auch billich geehrt/
Jetzt handt die weybschen geuch geleeert
Vnnd schaben all tag jr zwilchbacken
Sie


So groß gewalt vff erd nye kam
Der nitt zů zytten/end ouch nam
Wann jm syn zyl/vnd stündlin kam
von end des gewalttes
Noch fyndt man narren manigfalt
Die sich verlont vff jren gwalt
Als ob er ewiklich solt ston
Der doch dut/wie der schne zergon


Fueillet
De trop parler.
Qui scet sa langue refrener
Et de trop parler sa restraindre
Ne se voit a mal proferner
Par tristesse que lon doit craindre
Mais qui parletrop sans se faindre
Se deshonnore con fait la pie
Par trop garruler quon lespie

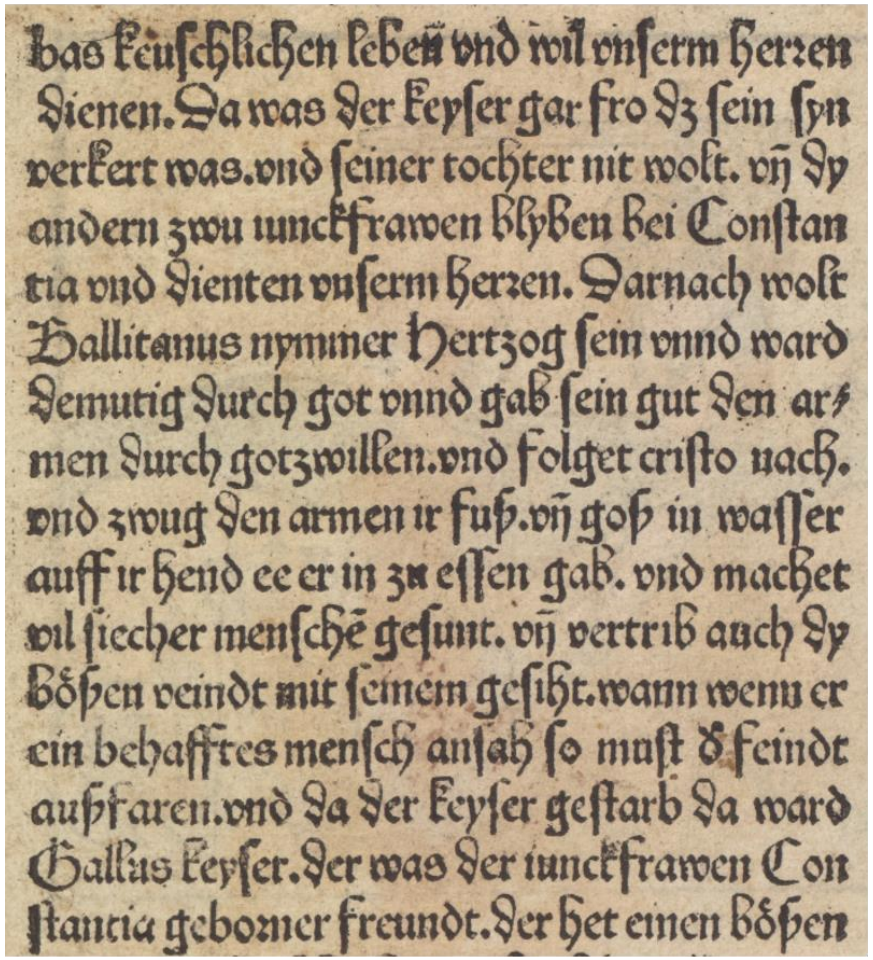


Frühdrucke bis 1600 – Ergebnisse

	Unerfahrene Nutzer	Erfahrene Nutzer
Erreichte CER	0,47% ± 0,22%	0,49% ± 0,30%
Transkribiertes Trainingsmaterial	988 Zeilen	927 Zeilen
Korrekturzeit pro Zeile	10s ± 5,2s	5,5s ± 2,4s
Segmentierungszeit pro Seite	1,1min ± 0,5min	0,6min ± 0,2min

*Jeweils Durchschnittswerte, ggf. mit Standardabweichung

Beispielergebnis mit ca. 0,6% CER



bas ksuſchlichen lcben vnd wil vnſerm herꝛen
dienen. Da was der keyſer gar fro dz ſein ſyn
verkert was. vnd ſeiner tochter nit wolt. vñ dy
andern zwu iunckfrawen blyben bei Conſtan
tia vnd dienten vnſerm herꝛen. Darnach wolt
Ballitanus nymmer Hertzog ſein vnnd ward
demutig durch got vnnd gab ſein gut den ar=
men durch gotzwillen. vnd folget criſto nach.
vnd zwug den armen ir fuß. vñ goß in waſſer
auff ir hend ee er in zu eſſen gab. vnd machet
vil ſiecher menſchē geſunt. vñ vertrib auch dy
bößen veindt mit ſeinem geſiht. wann wenn er
ein behafftes menſch anſah ſo muſt ð feindt
außtaren. vnd da der keyſer geſtarb da ward
tGallus keyſer. der was der iunckfrawen Con
ſtantia geboꝛner freundt. der het einen bößen

Gliederung

1. Einleitung
2. Submodule
3. Workflow
4. Live Demo
5. Evaluation
6. **Diskussion und Ausblick**

Zusammenfassung

- OCR4all vielseitig einsetzbar
  - Erfolgreiche Verarbeitung von (historischen) Drucken des 15. bis 21. Jahrhunderts
  - Hauptanwendung: lokale Installation beim Nutzer
  - Nutzung als Serveranwendung prinzipiell bereits jetzt möglich
- Ergebnisse und dafür notwendiger Aufwand stark abhängig …
  - … vom Material
    - Automatische Segmentierung möglich?
    - Passendes gemischtes Modell vorhanden?
    - …
  - … von den Nutzeranforderungen
    - Grad der semantischen Auszeichnung?
    - Ansprüche hinsichtlich Genauigkeit?
    - …

OCR-D  OCR4all

- Kooperationsvereinbarung Sommer 2020
  - Umsetzung von OCR-D Spezifikationen und Schnittstellen in OCR4all zum beiderseitigen Vorteil:
    - für OCR-D: bei Bedarf vereinfachter Zugang für die Nutzer, größere Reichweitere
    - für OCR4all: erweiterte Auswahl an Werkzeugen, Flexibilität
  - Fortlaufender Austausch (Schnittstellen, Skalierbarkeit, kommende OCR Entwicklungen, GT, …)
- Erfolgreicher Projektantrag *OCR4all-libraries* in dritter OCR-D Phase:
  - Kooperation zwischen GEI Braunschweig und Uni Würzburg (HCI, ZPD)
  - Volle Unterstützung der OCR-D Lösungen sowie deren Steuerung und Konfiguration über die GUI
  - Ermöglichung einer Massenverarbeitung von Werken bzw. Werkclustern
  - Ausbau von LAREX als visuelle Erklärungskomponente (Fehleranalyse, Vergleich von Workflows, …)
  - Optimierung der Usability
  - …

Verarbeitung von Handschriften

- Erfassung von Drucken und Handschriften konzeptionell sehr ähnlich
- OCR4all bereits jetzt vielseitig im Handschrifteneinsatz
- Systematische Evaluation in Kooperation mit Dr. Stefan Tomasek (Lehrstuhl für Ältere Deutsche Literatur, Uni Würzburg) anhand des Projekts „Konrad von Fußesbrunnen: Kindheit Jesu“
- Größtes Desiderat: robuste automatische Segmentierung
- Zeitnahe Anbindung vielversprechender, Baseline-basierter, trainierbarer (!) Lösungen:
  - Kraken (Kiessling: A Modular Region and Text Line Layout Analysis System)
  - Entwicklungen am Würzburger Lehrstuhl für Künstliche Intelligenz (Fischer, Hartelt, Gehrke, Puppe)
  - …

Aktuelle Entwicklungen und Planungen

- Zeitnah: weitreichend überarbeitete Version mit zahlreichen Verbesserungen hinsichtlich Stabilität und Flexibilität
- Mittelfristig: größere Flexibilität hinsichtlich Material und Anwendungsszenarien
  - Handschriften
  - Massenvolltextdigitalisierung
- Weiterer DFG Antrag in Vorbereitung, u. a. Fokus auf
  - Optimierung des kollaborativen Arbeitens (Projekt-, Nutzer- und Taskverwaltung, Ressourcenmanagement, Backup und Versionierung, …)
  - Bei Bedarf Steuerung einzelner Prozesse und Workflows aus LAREX heraus
  - Konfidenzbasierte Qualitätsanalyse und interaktive Nachkorrektur
  - Usability (Nutzerstudien, Überarbeitung des Schulungskonzepts und der Anleitungen, …)

Interesse an OCR4all?

- Probieren Sie OCR4all auf Ihren eigenem Rechner mit Ihrem Material aus
  - www.ocr4all.de
  - https://github.com/OCR4all/getting_started
- Wir helfen gerne bei Fragen und Problemen: ocr4all@uni-wuerzburg.de
  - Installation und Nutzung
  - Bug Fixes und Feature Requests
  - Projekt-spezifisches Consulting
  - Server Support, falls die eigene Hardware nicht ausreicht
  - …
- Wir sind auf Ihr (unverblümtes) Feedback angewiesen!
- … aber bedenken Sie bitte immer: OCR4all ist umsonst und Work in Progress ☺

Vielen Dank für Ihre Aufmerksamkeit!

Acknowledgements:

- **OCR4all Web App**: Dr. Herbert Baier-Saip, Maximilian Nöth, Dennis Christ, Alexander Hartelt, Nico Balbach, Kevin Chadbourne
- **LAREX Web App**: Maximilian Nöth, Kevin Chadbourne, Nico Balbach
- **Calamari**: Dr. Christoph Wick, Andreas Büttner
- **Tests, Anleitungen, Nutzer Support und Artwork**: Maximilian Wehner, Raphaelle Jung, …
- **Distribution via Docker und VirtualBox**: Björn Eyeselein, Yannik Herbst
- **Ideen und Feedback**: Dr. Uwe Springmann, Maximilian Wehner, Prof. Dr. Frank Puppe, Christine Grundig, Prof. Dr. Brigitte Burrichter, Prof. Dr. Joachim Hamm, …
- **Förderung**: DFG Förderinitiative „OCR-D“ Phase 2 und 3, BMBF Projekt „Kallimachos”, Lehrstuhl für Künstliche Intelligenz (Prof. Dr. Frank Puppe) und Zentrum für Philologie und Digitalität der Uni Würzburg