

Digitalisierung und Erschließung historischer Schulbücher mit Schwerpunkt religiöse Bildung – zwischen Masse und gezielten manuellen Eingriffen

Christian Reul

Forschungsstelle für digitale Anwendungen
in den Geisteswissenschaften (DACHS)

Universität Würzburg

27.11.2024



Agenda

1. Allgemeine Projektinfos

2. Genereller (OCR) Workflow

3. Layoutanalyse

4. OCR-Qualitätssicherung und –Optimierung

5. NER und NEL

Allgemeine Projektinfos

Digitalisierung und Erschließung historischer Schulbücher mit Schwerpunkt religiöse Bildung und Aufbau einer Wissensbasis für die bildungshistorische Forschung

- [DFG-Projekt](#), Förderlinie „Digitalisierung und Erschließung“
- Kooperation mit Dr. Anke Hertling (Leibniz-Institut für Bildungsmedien | GEI)
- Laufzeit drei Jahre, Start 07/24
- Ziele:
 - Erfassung von deutschspr. Religionsschulbücher und schulischen (Erst-)Lesebüchern
 - Knapp 2.000 Bände mit fast 500.000 Seiten
 - Zeitraum 1618 bis 1870
 - Systematische Erfassung von Akteuren aus der historischen Schulbuchproduktion unter Einsatz von Named Entity Recognition (NER) und Named Entity Linking (NEL)

Vorrede.

ser, oder einer Geschichte, verschiedene Dinge zu merken, so sind zwar solche in der Ordnung nach einander hergesetzt; Dem Lernenden aber kann die Sache erleichtert werden, wann man durch kleinere Fragen Dieselbe annoch zergliedert, und die Antwort solcher gestalt von ihm heraus locket; 3. E. p. 60. werden Caligula Laster und sonderlich seine Verschwendung beschrieben, weil diese nun, nach einander zu erzehlen, einem Anfänger in der Historie schwer fallen möchten, so kann ad §. 1. gefraget werden: Wie viel Geld hat Caligula durchgebracht? ad §. 2. Wie ging er mit seinen Schwestern um? ad §. 3. & 4. Wie tractirte er sein Pferd? und so an andern Orten mehr.

Was die gemengte oder umgekehrte Fragen pag. 22. 29. 36. 45. 2c. denen Lernenden für einen Vorteil und Impression geben, wird man bei dem ersten Anblick derselben alsobald urtheilen,

22 MONARCHIÆ PRIMÆ

8. Aeneas succediret setnem Schwieger-Vater Latino im Regiment.

9. Er ist der erste König der Latiner von denen/die nach der Zerstörung Troja in Italien kommen.

10. Von dem Todt Aeneas bis auff die Erbauung der Stadt Rom werden 426. Jahr gezehlet.

CENTURIÆ TERTIÆ. Decas I.

1.

Thineus XXX.

Thineus regieret 30. Jahr / von An. M. 2831.

2. Die Labe des Bundes wird von den Philistern weggenommen. 1. Sam. 4. v. 11.

3. Samuel succediret dem Eli An. M. 2850. und ist so wohl ein Prophet als ein Politischer Richter gewesen.

4. Dem Ascanio Aeneas Sohn folget im Regiment nach Sylvius, von welchem hernach die Successores Sylvii genennet werden / weil er in einem Walde gebohren.

5. Die Heraclidæ kommen in den Pelopponesum, und nehmen Spartam, Argos und Messenen ein/welches geschehen 80. Jahr nach der Verwüstung Troja.

Dercylus XXXI.

6. Dercylus regieret 40. Jahr von A. M. 2861.

7. In

Eroberungspläne zu fördern, und ließ Truppen in Polen einrücken. Um nicht das ganze Land eine Beute Rußlands werden zu lassen, schloß Friedrich Wilhelm II., wie sein großer Vorgänger, mit der Kaiserin einen Teilungsvertrag, wonach große Stücke des östlichen Polens an Rußland, Danzig, Thorn und die von nun an als Südpreußen bezeichneten Gebiete an Preußen fielen. Osterreich war an dieser Teilung nicht beteiligt; um so eifersüchtiger beobachtete es die Vergrößerung Preußens.

Quelle polnische Teilung. 1793.

Jetzt entstand in Polen eine starke nationale Bewegung zum Schutze des zerstückelten und in seinem Dasein bedrohten Vaterlandes; an ihre Spitze trat Kosciuszko. Aber trotz verzweifelten Widerstandes erlagen sie dem russischen General Suworow, der Praga, die Vorstadt Warschau, erstürmte und die Hauptstadt darauf einnahm. Nunmehr folgte die dritte Teilung Polens, das aufhörte ein selbständiger Staat zu sein. Preußen erhielt Neupreußen mit Warschau, Osterreich Westgalizien, Rußland die großen Gebiete, die übrig blieben.

Dritte Teilung 1795.

§ 10. Der erste Koalitionskrieg. Bonaparte. Die wichtigsten Schauplätze des ersten Koalitionskrieges waren Belgien, die Rheinlande und Oberitalien. Der Organisator der französischen Armeen war Carnot, welcher Mitglied des Wohlfahrtsausschusses war. Sie wurden auf Grund einer allgemeinen Aushebung gebildet, die unter den jungen Leuten vom 18. bis zum 25. Jahre stattfand; so brachte die französische Revolution wieder den Grundsatz der allgemeinen Wehrpflicht zur Geltung, während die Heere der Verbündeten aus Berufs Soldaten bestanden. Ohne sich an die Regeln der überlieferten Kriegskunst zu lehnen, gingen die kühnen, jungen Generale der Franzosen frisch auf den Feind los. Verluste konnten sie durch neue Aushebungen leicht ersetzen; ihren Unterhalt beschafften sie sich durch Requisitionen. Bald waren sie fast überall im Vorteil. Das linke Rheinufer mußte von den Verbündeten geräumt werden; ebenso eroberten die Franzosen Holland und verwandelten diesen Staat in eine batavische Republik.

1793 bis 1797.

Die Kriegsführung.

Da entschloß sich Friedrich Wilhelm II., teils weil ihm die Geldmittel zur Fortführung des Krieges fehlten, teils weil seine Truppen durch die polnischen Ereignisse in Anspruch genommen waren, zum Frieden mit Frankreich, der zu Basel im Jahre 1795 abgeschlossen wurde. Durch eine Demarkationslinie wurde Norddeutschland von dem süddeutschen Kriegsschauplatz geschieden; dort herrschte Friede, hier tobte der Krieg weiter. Während die Waffen über das Schickal Europas entschieden, be-

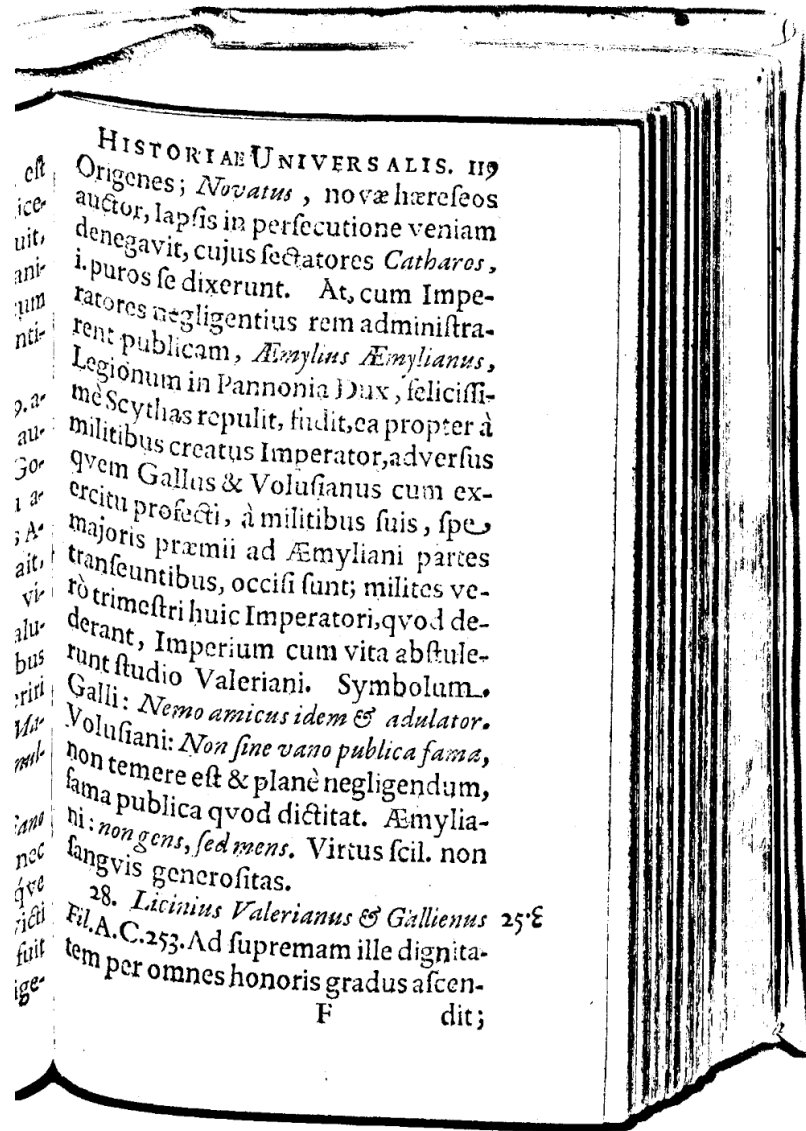
Eroberung Hollands.

Friede von Basel. 1795.

326 folgte deshalb ein 22jähriger Krieg gegen die Samniter (326 bis 304, zu der Zeit, als eben Alexander das Perserreich zerstört hatte; vergl. Liv. IX, 17), in welchem die Römer unter manchen Mißgeschicken ihre ganze Heldenthat entwickelten. Nachdem der große samnitische Feldherr Pontius ein römisches Heer in den Engpässen von Caudium eingesperrt hatte (321), mußte dasselbe schimpflich unter dem Joch hingehen; als der Senat den Consul, welcher den Vertrag eingegangen war, auslieferte, wies der ehrenhafte Pontius diesen zurück. Der altpatricische Papius Cursor rückte dann die Römer und wurde der Hauptheld dieses Krieges. — Auch die Etrusker und Umlrer hatten sich inzwischen den Samnitem angeschlossen; doch siegte Rom „durch seine centrale Lage und die Festigkeit seines Staatsverbandes“, und im Frieden erkannte Samnium die Oberhoheit auf der Römer an.

Die Plebejer und Patricier, die unter den schweren Kämpfen immer verbunden waren, beendeten jetzt ihren Zwist um die Zulassung zu den höheren Ämtern, indem den Plebejern endlich auch ein gleicher Antheil an dem **Priesterthume** (ogulnisches Geseß) gewährt wurde (300). Schon während des Krieges war die persönliche Schuldhaft aufgehoben. — Ein Versuch des Appius Claudius Censor (der sich durch Anlage der via Appia von Rom nach Capua und der ersten Freileitung Volksgunst erwarb), die nicht grundsätzliche Klasse (auch Freigelassene) willkürlich zum Stimmrecht heranzuziehen, um ein patricisches Regiment auf den Vöbel zu stützen, wurde dadurch unschädlich gemacht, daß D. Fabius (hier von Maximus genannt) diese neuen Bürger auf die vier städtischen Tribus einschränkte, wodurch ihre Stimmen von denen der alten (ackerbauenden) Plebejer in den ländlichen Tribus, deren Zahl bis 241 allmählich auf 31 stieg, überwogen wurden.

d. Dritter (zehnjähriger) Samniterkrieg. Wegen die weitere Ausbreitung der römischen Herrschaft, die immer mehr durch starke Militär-Colonien gesichert wurde, traten (298) noch einmal die Völker Italiens zusammen, als sich neue gallische Einwanderer den Etruskern zu Söldnern anboten und diese wie die Umlrer sich mit den Samnitem verbündeten. Als die letzteren von Umbrien (im Osten der Apenninen) gegen die Römer ziehen wollten, siegten diese unter D. Fabius Maximus bei Sentinum, indem sich Decius Mus der Sohn dem Tode weihete. Nachdem endlich der edle Pontius gefangen und schmachlich hingerichtet war, mußten die Samniter, Etrusker und Umlrer die Oberherrlichkeit der Römer anerkennen. M. Curius Dentatus, der den Frieden mit den Samnitem erkämpfte, unterwarf auch noch die Sabiner; dann brachte er für die verarmten Plebejer ein Ackergesetz in Vorschlag, das große Kämpfe hervorrief. Nach einer nochmaligen Auswanderung der Plebs (auf den Janiculum) setzte aber der Dictator Hortensius nicht nur Schuldenerlaß und Ackervertheilung durch, sondern nach einer Rogation desselben



HISTORIAE UNIVERSALIS. II 9
 Origenes; *Novatus, novae haeeseos auctor, lapsus in persecutione veniam denegavit, cujus sectatores Catharos, i. puros se dixerunt. At, cum Imperatores negligentius rem administrarent publicam, Aemilius Aemilianus, Legionum in Pannonia Dux, felicissimè Scythas repulit, fudit, ea propter à militibus creatus Imperator, adversus quem Gallus & Volusianus cum exercitu profecti, à militibus suis, spe majoris praemii ad Aemiliani partes transeuntibus, occisi sunt; milites vero trimestri huic Imperatori, quod deriderant, Imperium cum vita abstulerunt studio Valeriani. Symbolum. Galli: Nemo amicus idem & adulator. Volusiani: Non sine vano publica fama, non temere est & planè negligendum, fama publica quod dicitur. Aemilianii: non gens, sed mens. Virtus scilicet non sanguis generositas.*
 28. *Licinius Valerianus & Gallienus* 255
 Fil. A. C. 253. *Ad supremam ille dignitatem per omnes honoris gradus ascendit;*

est
 ice
 uit
 an
 um
 nti
 2. a
 au
 300
 i a
 ait
 vi
 alu
 bus
 riri
 Ma
 mil
 iam
 nec
 que
 rici
 fuit
 lige

DES XVII. SECLI DER OESTERREICHISCHEN KAISER. TAB.

DEUTSCHE KAISER. ...
 I. OBERÖSTERREICHISCHER KREIS. ...
 II. NIEDERÖSTERREICHISCHER KREIS. ...
 III. WESTPHÄLISCHER KREIS. ...
 IV. NIEDERRHEINISCHER KREIS. ...
 V. OBERRHEINISCHER KREIS. ...
 VI. SAARLÄNDISCHER KREIS. ...
 VII. SCHWÄBISCHER KREIS. ...
 VIII. RHEINISCHER KREIS. ...
 IX. OESTERREICHISCHER KREIS. ...
 X. BAHYENISCHER KREIS. ...

DES DEUTSCHEN REICHS STÄNDE.
 I. OBERÖSTERREICHISCHER KREIS. ...
 II. NIEDERÖSTERREICHISCHER KREIS. ...
 III. WESTPHÄLISCHER KREIS. ...
 IV. NIEDERRHEINISCHER KREIS. ...
 V. OBERRHEINISCHER KREIS. ...
 VI. SAARLÄNDISCHER KREIS. ...
 VII. SCHWÄBISCHER KREIS. ...
 VIII. RHEINISCHER KREIS. ...
 IX. OESTERREICHISCHER KREIS. ...
 X. BAHYENISCHER KREIS. ...

Anmerkungen.
 Der Reichstage wurde gehalten zu Regensburg, ...

Agenda

1. Allgemeine Projektinfos
- 2. Genereller (OCR) Workflow**
3. Layoutanalyse
4. OCR-Qualitätssicherung und –Optimierung
5. NER und NEL

Genereller (OCR) Workflow

- Binarisierung/Preprocessing mit SBB-Binarize und OCRopus
- Layoutanalyse → nächster Abschnitt
- Erkennung voraussichtlich mit Calamari, vorab allerdings noch Tests; Neue gemischte Modelle werden trainiert/evaluiert
- OCR-Qualitätssicherung und -Optimierung → übernächster Abschnitt
- NER und NEL → überübernächster Abschnitt

Agenda

1. Allgemeine Projektinfos
2. Genereller (OCR) Workflow
- 3. Layoutanalyse**
4. OCR-Qualitätssicherung und –Optimierung
5. NER und NEL

Überblick

- Ursprünglicher Plan: OCR-D Workflow Suche
- Aber: lieber direkt voll auf trainierbare Methoden setzen
 - Besser und nachhaltiger
 - Paralleles [Projekt zur Layoutanalyse](#) (mit Stabi Berlin und SLUB Dresden) → Synergieeffekte
 - aktuell Fokus auf Kraken, später ggf. Eynollah u. a.
 - gezielt Trainingsdaten erstellen, Active Learning
- Aktuell viel ausprobieren, Erfahrungen sammeln
- Erste Experimente
 - Pragmatische Regionentypen
 - Recht differenzierte Auszeichnung
 - Für Training vorerst stark eingedampft (Graphic Region, paragraph, marginalia)
 - Roher Output, kein zusätzliches Postprocessing

Beispiele – Trainingsdaten I

paragraph

— 135 —

Bald sieht man von ferne die Türme der Paläste und Tempel der Hauptstadt, in der der Pharao Hof hält. Freilich, zuerst gilt es die äußeren Bezirke der Stadt zu durchheilen und da ist nicht viel Schönes zu sehen: kleine, armselige Lehmhütten, dazwischen unregelmäßige, schmale Gäßchen oder schmutzige Plätze, oft auch ein Tümpel, in dem Ochsen baden. Das niedere Volk ist sehr arm, es fristet nur eben sein Dasein und kann seine Hütten offen stehen lassen, da es dort nichts zu stehlen gibt.

Vorstadt.

Kommt man dagegen näher zum Mittelpunkt der Stadt, so wird es allmählich anders. Die Häuser sind größer und manchmal recht hoch, so daß man in den engen Gassen kaum den Himmel sieht. Dann öffnet sich ein Platz, auf dem ein Markt abgehalten wird. Die Leute drängen sich eifrig handelnd und fellschend durcheinander und haben für nichts anderes Augen und Ohren. Und der Handel ist auch manchmal gar nicht einfach. Eigentliches Geld gibt es nicht, sondern nur Edelmetallstücke, die bei jedem Handel neu abgewogen werden müssen, wobei man immer Angst hat, übervorteilt zu werden; daher ziehen viele den Tausch vor. Hier bietet der eine selbstverfertigte, dauerhafte Sandalen zum Tausch für einen Topf mit wohlriechender Salbe, dort werden Perlenstränge gegen Weihrauch und andere Wohlgerüche verhandelt. An einer anderen Stelle wird um einen Stier gefeilscht und man muß für ihn 1 Matte, 5 Maß Honig, 8 Maß Öl und noch 7 kleinere Dinge geben. Das muß man alles genau in Metallwert ausrechnen und dann erst kann der Tausch durchgeführt werden. Vom Markte gehen enge Gassen nach verschiedenen Seiten aus. Hier wohnen Kaufleute und Handwerker: Schuster, Bäcker und Zuckerbäcker, Garküche, Goldschmiede, Tischler; aber auch Wirtstuben gibt es, in denen Wein und ein Gerstengebräu, unserem Bier ähnlich, geschänkt wird. Auch Waren aus den asiatischen Kulturländern werden hier verkauft und ein gar buntes Treiben herrscht, fast wie auf dem Markte.

Die Stadt.

Handel.

marginalia

Gewerbetreibende.

Nun weiten sich die Gassen und nach einiger Zeit stehen wir vor der königlichen Hofburg, die sich fast wie eine Stadt in der Stadt ausnimmt. — Im Augenblicke, da wir uns nähern, tritt die Spitze eines Zuges aus dem Palast hervor. Voraus eilen Läufer, die die Aufgabe haben, das Herannahen des Pharaos zu melden. Alles Volk weicht sehen zur Seite, es öffnet sich ein weiter freier Raum. Nun folgt der Zug: an der Spitze eine Abteilung auserwählter Leibgarden, dann Fahnenträger, andere Leute, die große Wedel, Zeichen der königlichen Würde, tragen, dann Krieger, die mit Lanzen, Beil und Schild gerüstet sind, und endlich der Pharaos selbst. Allein steht er auf seinem Wagen, in ein langes, wallendes Gewand gehüllt; wie ein Gott erscheint er dem Volke, das sich vor ihm auf den Boden wirft. Hinter ihm kommt die Königin, gleichfalls im wallenden Gewand, und dann die Prinzen und hohen Beamten. Der König wird von Sklaven geleitet, die große Wedel über sein Haupt halten. So zieht er in den Tempel, um dort seinem Gott, als dessen Sohn er sich betrachtet, ein Opfer darzubringen.

Hofburg.
Auszug des Pharaos.

Separator
Region

page number

— 187 —

Ist will i dein Nachfolger sein.
Herr! machs nach de'm Bertanga.

3. Ein Bilgram bin i halt monneh,
Muß rafa fremde Strofa,
Das bitt i di, mein Gott und Herr
Du wirst mi nit verlofa

4. Den Glauba hob i frei befannt,
Das dorf i mi nit schäma,
Wenn mo mi glei ein Kezer neunt
Und tu mir's Leba nehma.

5. Muß i glei in das Stend fort,
Wil i mi do nit wehra,
So hoff i do, Gott wird mir dort
Och gute Fründ Leichera

6. Herr, wie du wilt, so gib mi drein,
Bei dir wil i verbleiba,
I will mi gerr dem Wille dein
G'duldig unerschreiba.

7. Mueß i glei fort, in Gottes Nam,
Und wird mir alles genomma,
So was i wol, die Himmelskron
Wer i emochl bekomma

8. So mueß i hent von meinem Haus,
Die Kindel mueß i laffa,
Mein Gott, es treiß mir Zähref aus
Zu wandern fremde Stroja

9. Mein Gott, fähr mi in eine Stadt,
Wo i dein Wort kam heba,
Darin will i die früh und zw
Zu meinem Herzal loba.

10. Sol i in diesem Jammeral
Noch länger in Bruut leba,
So hoff i da, Gott wird mir dort
Ein begre Wohnung geba.

52. Türken und Franzosen

Ziegler: Deutsche Soldaten-Kriegslieder aus 5 Jahrhunderten
(1380—1871) Leipzig, 1894.
Dr. W. Schilling: Liederbuch, S. 219/20

Lirt, jetzt ist es dahin kommen,
Daß du werdest ganz lathet,
Dann wir haben's vorgenommen,
Daß dein Reich zu Ende geht.

Gwar der arge Schandfranzose,
Der die Welt außkänferit,

Image Region

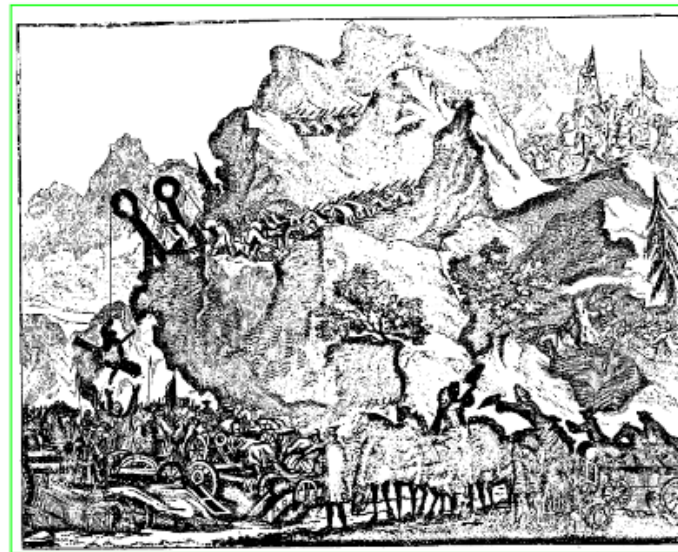


Fig. 43. Der „sehr mühsame und wunderswürdige Marsch“ der kaiserlichen Armee über die Tirolischen und Alt-Norischen Alpen im Jahre 1791.

Anschläge der Gegner zu vereiteln. Ludwig XIV., mit seinen bisherigen Feldherren unzufrieden, übergab den Oberbefehl an Villeroy (sprich: Willroa), der versprochen hatte, er werde Eugen tot oder lebendig dem Könige überliefern. Aber Eugen überrumpelte den Prahler zur Nachtzeit in der Mitte seiner Truppen und führte ihn, ehe die Seinigen ihm Hilfe bringen konnten, als Gefangenen fort. Selbst in Paris sang man Spottlieder auf Villeroy. Auch gegen Villeroy's Nachfolger Vendôme (sprich: Wandohn)

1) Südöstl. von Verona; nicht zu verwechseln mit der Stadt Carpi südlich vom Po.

caption

heading

other

footnote

Beispiele – Trainingsdaten II

Rückzug der Skythen.
 Rückzug der Perser.
 Die Griechen bei der Brücke.



Fig. 8. Darius mit Dienern (über ihm Ormuzd). Wandbild im Palast zu Persepolis.

wichen zurück und ließen sich in keine offene Schlacht ein, sondern verwüsteten nur das Land und verschütteten die Brunnen, bis den Persern die Lebensmittel ausgingen und Darius den Befehl zum Rückzug gab. Jetzt begannen die Skythen die Verfolgung, umschwärmten das Heer von allen Seiten und bedrängten es so sehr, daß es froh sein mußte, wieder den Ister zu erreichen. Zum Glück fand man dort die Brücken noch vor; der Mann, dem Darius seine Rettung verdankte, war Histäus, der Stadtherrscher (Tyrann) von Milet, einer der griechischen Städte an der Westküste Kleinasiens. Es waren nämlich Skythen an den Fluß gekommen und hatten den Griechen geraten, die Brücken abzubauen, dann würde das ganze persische Heer zugrunde gehen; Miltiades, ein Griechenfürst vom Hellespont, hatte dem beigestimmt, da dann die Griechen Kleinasiens sich vom persischen Joche befreien könnten, aber Histäus widersprach und meinte umgekehrt, die Stadtherrscher hätten alle Ursache, die Rettung des persischen Königs und Heeres zu wünschen, da sie nur durch des Königs Macht in ihrer Stellung geschützt würden. Ohne diesen Schutz hätten die Bürgerschaften ihrer Städte sie schon längst vertrieben. So blieben die Brücken erhalten

und Darius war gerettet. Histäus wurde belohnt, Miltiades floh in seine griechische Heimat.

nun wenigstens den Übergang über die Etsch zu verhindern, zersplitterte aber, da Eugen ihn über die Stelle, wo der Übergang beabsichtigt war, geschickt in Ungewißheit zu erhalten wußte, seine Truppen von den Alpen bis zum Po. Als daher Eugen die Etsch bei Carpi¹⁾ wirklich überschritt, fand er nur wenig zahlreiche feindliche Truppen, die er schlug. Unaufhaltsam drang er dann in der Lombardei vor. Seine kühnen Reiter, die weithin das Land durchstreiften, fingen die Berichte der französischen Generale an ihren König und die königlichen Befehle an die Generale ab, und so wußte Eugen trotz seiner geringen Truppenzahl alle

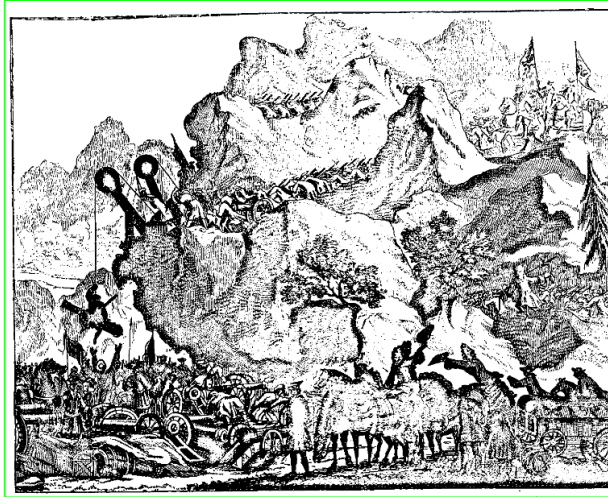


Fig. 43. Der „sehr mühsame und wunderswürdige Marsch“ der kaiserlichen Armee über die Tirolischen und Alt-Norischen Alpen im Jahre 1701.

Anschläge der Gegner zu vereiteln. Ludwig XIV., mit seinen bisherigen Feldherren unzufrieden, übergab den Oberbefehl an Villeroy (sprich: Willroa), der versprochen hatte, er werde Eugen tot oder lebendig dem Könige überliefern. Aber Eugen überrumpelte den Praher zur Nachtzeit in der Mitte seiner Truppen und führte ihn, ehe die Seinigen ihm Hilfe bringen konnten, als Gefangenen fort. Selbst in Paris sang man Spottlieder auf Villeroy. Auch gegen Villeroy's Nachfolger Vendôme (sprich: Wandohm)

¹⁾ Südöstl. von Verona; nicht zu verwechseln mit der Stadt Carpi südlich vom Po.

nahe stand, doch von ihnen stets als ein barbarisches angesehen wurde. Die wenigen Städte waren meist griechische Kolonien, die Handel und Verkehr im Lande beherrschten.

Philipp II. In den Perserkriegen hatten die Mazedonierkönige öfters die Vermittler zwischen den Persern, deren Oberhoheit sie damals anerkannten, und den Griechen gespielt; sonst waren sie wenig hervorgetreten. Pelopidas hatte, wie oben erwähnt, Thebens Einfluß auch bis nach Mazedonien ausgedehnt und den jungen Königssohn Philippus als Geisel nach Theben mitgebracht; nach mehrjährigem Aufenthalt daselbst kehrte der Prinz in seine Heimat zurück und gelangte im Jahre 359 zur Regierung.

Philipp hatte die Zeit in der Fremde gut benützt: er kannte die Griechen und ihre Schwächen sehr genau und faßte den kühnen Plan, sich nach und nach der nächstgelegenen Kolonien zu bemächtigen, seine Macht allmählich auszudehnen und dann mit einem Schlage ganz Hellas unter seine Oberhoheit zu bringen. Er war wohl der Mann dazu, das alles auszuführen: ebenso tapfer und entschlossen als gewandt und schlaue, scheute er kein Mittel, um sein Ziel zu erreichen; in allen Griechenstädten gewann er sich Freunde, größtenteils durch Bestechung; denn, wie er sagte, „einem mit Gold beladenen Esel widersteht keine Festung“. Aus seinen Mazedoniern schuf er sich ein ausgezeichnetes Heer; der Adel, „die Freunde des Königs“, bildete die gut bewaffnete Reiterei und den Kern des Fußvolks, aber auch die freien Bauern dienten eine Zeitlang im Heere als Hopliten oder leichter bewaffnete Fußsoldaten. Sie stellten sich im Kampfe dichtgeschlossen auf, Schild an Schild gedrängt, 16 Glieder tief; die ungewöhnlich langen Lanzen (6 m!) ermöglichten es, daß die fünf vordersten Glieder wirklich zum Kampfe kamen, indes die andern nachdrängten; diese mazedonische „Phalanx“ (Schlachtreihe) galt durch lange Jahre als unüberwindlich.

Philipp und Athen. Bald hatte Philipp unter verschiedenen Vorwänden die bedeutendsten griechischen Kolonien an den Grenzen Mazedoniens eingenommen, ohne ernstlichen Widerstand zu finden. Die Athener, die bisher in diesen Gebieten den Haupteinfluß gehabt hatten, ließen ihn ruhig gewähren; die Stadt war längst nicht mehr, was sie einst gewesen. Die Bürger hatten sich allmählich entwöhnt vom Kriegsdienst, den sie gemieteten Söldnern und Feldhauptleuten überließen. Wie in jener Zeit überhaupt viele

Philipp in Theben. Seine Thronbesteigung, 359 v. Chr.

Philipps Pläne.

Sein Charakter.

Das Heer.

Athens Schwäche.

Söldnerwesen.

Beispiele – Ergebnisse I

128

Das Mittelalter.

Keine Straßenbeleuchtung.
Abwechslung.

Abend und in der Nacht, wo es keine Laternen gab, und jeder, der das Haus verließ, ein kleines Licht bei sich führte, aufrecht zu erhalten.

Draußen auf dem Lande und hoch oben in den Felsenestern verlief das Leben im ganzen sehr eintönig. Der Burgherr vertrieb sich die freie Zeit gern mit der Jagd und Beschlagen. Mancher zog auch hinaus in die Welt, um für seine „Herrin“ Abenteuer zu bestehen. Dazu kamen die Hoffahrt und die Turniere. Die „Dörpel“ oder „Dörpschen“ (Dörfer, Bauern) erfreuten sich am Tanz unter der Linde; in den Städten aber entstanden die Feste der wehrhaften Bürgerschaft, die Schützenfeste. Das alles geschah zur hochgelobten Sommerszeit oder im Frühling. In den langen, einsamen Winternächten aber, die, weil man der kalten Luft wegen gewöhnlich die Fenster zugstopfte, auch den kurzen Tag mit in sich wuschen, war es ein Labfal für jung und alt, wenn Fremde kamen, die etwas zu erzählen wußten, oder Spielleute, die außerdem neue Tanzweisen vortrugen und alte Heldensieder zu singen und zu sagen verstanden. Dies waren „Fahrende“, die man wohl gering schätzte, die aber niemand mißhen mochte. Reisen war, falls nicht ein kraftvoller König den Landfrieden aufrecht erhielt, eine gefährvolle und bei der schlechten Beschaffenheit der Wege sehr unangenehme Sache. Aus letzterem Grunde reiste man zu Pferde.

Die Fahrenden

Reisen.

Baukunst.
Gotische Bauweise.

Die Künste. In Nordfrankreich war die sog. gotische Bauweise aufgetommen. In ihr, die in leicht aufstrebenden, „Lauben“ nachahmte, geöffneten Bauten wohl die Art der germanischen „Lauben“ nachahmte, wurden herrliche Kirchen, dazu Rat- und Kunsthäuser und Kaufhallen errichtet. — Das Eigentümliche der gotischen Bauweise war der Spitzbogen. Da nun spitzbogige (aus Kreissegmenten gebildete) Gewölbe einen geringeren Schub und Druck als die in Halbkreisbogen konstruierten romanischen Gewölbe ausübten, bedürfen sie nur einer Unterstützung der Endpunkte. Und zwar trugen den senkrechten Druck des Gewölbes Pfeiler, den schwachen Seitenschub hielten die den inneren Pfeilern entsprechenden, an der Außenseite angebrachten starken Mauerpfeiler (Strebpfeiler) auf, die, in der Höhe abgestumpft, mit einem Diefel oder Türmchen gekrönt waren. Damit war der gedrungene Mauerbau aufgelöst in ein reiches, kühn aufstrebendes Baugerippe. Die den Raum abschließenden Wände verwandelten sich in leichte Füllungen, und dieser Wandverschluß wurde durch hohe und weite Fenster durchbrochen.

Glieder.
Schmuck.

Für die Glieder und die Schmückung galt der Grundsatz, daß diese in enger Beziehung zum Aufbau stehen, jene das Wesen desselben wieder spiegeln und in unendlichen Abänderungen zur Darstellung bringen sollten. Die schlanken Pfeiler wurden zu Bündelpfeilern zusammengefaßt und reich gegliedert, sie bestanden aus einem Kern und Säulchen, „Dienste“ genannt, die um ihn herumgestellt waren. Das Kapital zeigte die Form des Kelches; zum Schmuck diente naturgetreu gebildetes Laubwerk. Das Kreuzgewölbe erschien als neß- oder sternförmiges Gewölbe. Die Fenster waren zu ausgedehnt, darum teilte man sie in Gruppen; den Raum unter dem Bogenscheitel¹⁾ füllte man mit geometrischen Gebilden, „Maßwerk“.

1) In der späteren Zeit der Gotik bildete man den Bogen in der Gestalt eines Gesträndes. Auch gab man da den Seitenrippen die Höhe der Mittelschiffe.

§ 48. Kalvin. Der Jesuitenorden. Die Kirchenversammlung von Trient. 195

„Konvikte“, in denen der Nachwuchs für die Prediger und Lehrer herangezogen wurde. Prächtige Schlösser¹⁾ verkündeten das kraftvolle Selbstgefühl und den steigenden Reichtum der Fürstengeschlechter.

Schlösserbauten
in deutscher
Renaissancezeit.

6. Karls Abdankung und Tod (1558). So war der stolze Plan, den der junge Kaiser Karl entworfen, gescheitert; weder eine habsburgische Universalmonarchie noch die dazu unerläßliche Vereinigung der Religion war trotz alles noch so eifrigen und geschickten Wirkens durchzusetzen gewesen. Verbittert über das Fehlschlagen seines Lebenswerkes, zog sich der Kaiser, überdies krank und früh gealtert, vom Schauplatz zurück. Nachdem er bereits im Jahre 1554 seinem Sohne Philipp Mailand und Neapel übertragen hatte, legte er in feierlicher Versammlung der burgundischen Stände auch die Regierung der Niederlande zu jenes Gunsten nieder (1555). Im nächsten Jahre verzichtete er auch auf Deutschland und Spanien, von denen jenes Ferdinand, dieses Philipp erhielt (1556). Ohne Mönch zu werden, weilte er von da ab in der Nähe des Klosters des heiligen Hieronymus bei San Juste im spanischen Estremadura, wo er im Jahre 1558 starb.

Erfolg des Wissens
Kaisers V.

Niederlegung
seiner Krone.

Philipp II. König
von Spanien, der
Niederlande und
den italienischen
Gebieten.

Ferdinand I.
Kaiser 1558.

Gegen Verpflichtung auf den Augsburger Religionsfrieden wurde in demselben Jahre (1558) Ferdinand zum römischen Kaiser gewählt.

B. Die Gegenreformation. Die Ausbildung der Fürstenmacht.

§ 48. Kalvin. Der Jesuitenorden. Die Kirchenversammlung von Trient.

Die deutschen Protestanten hätten alle Ursache gehabt, zusammenzuhalten und die unter ihnen bestehenden Meinungsverschiedenheiten nicht bis zur Spaltung oder gar zu politischem Gegensatz ausreifen zu lassen. Statt dessen aber verbrauchten sie ihre Kraft in innerem Haber, den auch eine Fürstenversammlung zu Raumburg a. d. Saale (1561) nicht zu bannen vermochte. Dies war um so verhängnisvoller, als bereits die alte Kirche alle ihre Kräfte zusammenfaßte, um nicht nur dem ferneren Ausbreiten der evangelischen Lehre zähen Widerstand entgegenzusetzen, sondern auch den Kampf um die Rückgewinnung des verlorenen Besitzes entschlossen aufzunehmen.

Der deutsche
Protestantismus
nach dem Jahre
1555.

Die alte Kirche.
Sammlung zum
Gegenangriff.

Genf.

Die Verteidigung und den Gegenangriff gegen die anhebende „Gegenreformation“ übernahmen auf protestantischer Seite die Reformierten kalvinistischer Richtung, deren Hochburg die Stadt Genf war.

1. Kalvin. Als es offenbar wurde, daß Savoyen darauf ausginge, die fast ganz selbständige Stadt Genf sich einzuverleiben, bildete sich hier eine Freiheitspartei, die Eidgenos oder Huguenots²⁾, welche den Anschluß an die Schweiz bewirkte.

Genf tritt zur
Schweiz über.

1) So in Dresden, München, Stuttgart, Heidelberg. Hier wurde der Ottheinrichsbau in den Jahren 1556–1563, der Friedrichsbau 1601–1607 errichtet.
2) François Hugues war ihr Vorkämpfer.

134 § 47. Renaissance der Wissenschaften und der Humanismus.

2. Der Humanismus in Deutschland. Während in Italien die Renaissance in erster Linie eine gesteigerte Pflege des Schönen und eine verfeinerte Lebensführung im Gefolge hatte, in ihrer Entartung aber auch den gefälligen Deckmantel leichtfertiger Sitten bot, gedieh sie in Deutschland zur strengen Schule der Geistesbildung und der Erziehung. Die Humanisten, wie man die Vertreter der neuen Bildungsrichtung nannte, erkämpften hier bald die Lehrstühle der Schulen und Universitäten. Auch neue Hochschulen entstanden in rascher Folge, so in Ingolstadt 1472 (gegründet vom bayerischen Herzog Ludwig dem Reichen von Landshut), in Trier 1473, in Mainz und in Tübingen 1476, in Wittenberg 1502, in Frankfurt a. d. Oder 1506.

Unter den älteren der **deutschen Humanisten** — sie gaben sich zumeist latinierte oder gräzisierte Namen — ragen hervor: Rudolf Agricola (in Heidelberg, † 1485), Konrad Celtes (in Ingolstadt und Wien, † 1508), Johann Neuklin (in Tübingen und Ingolstadt, † 1522), Desiderius Erasmus (in Basel, † 1536); außerdem Konrad Peutinger aus Augsburg († 1547), Willibald Pirheimer zu Nürnberg († 1530) und seine Schwester Charitas, Abtissin eines dortigen Klosters. Andere Humanisten, wie Ulrich von Hutten († 1523 in der Schweiz) und Philipp Melancthon († 1560 in Wittenberg), gehören als Streitgenossen Luthers der Reformationsperiode an.

§ 48. Renaissance der Künste. I. Baukunst.

1. In Italien. Durch die neue Zeitströmung begünstigt, gelangte seit der Mitte des 15. Jahrhunderts zunächst in Italien wieder die Architektur des Korinthisch-Römischen Baustils zur vorherrschenden Geltung. Im Gegenfag

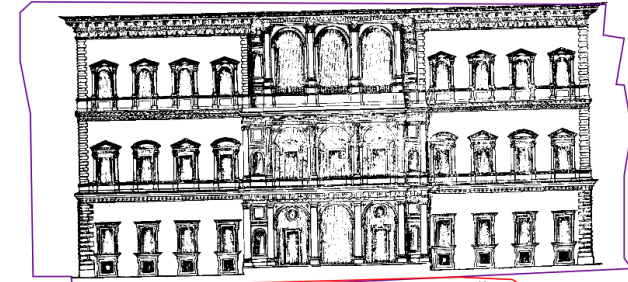


Fig. 11. Der Palazzo Borromeo in Rom (Renaissancezeit).

zur Gotik schufen die hochgebildeten Meister der italienischen Renaissance seltene Bauten mit ruhigen und harmonischen Grund- und Zierformen.

Beispiele – Ergebnisse II



30. Stifterin. Dom zu Naumburg.



31. Grabmal Heinrichs d. L. Braunschweig.

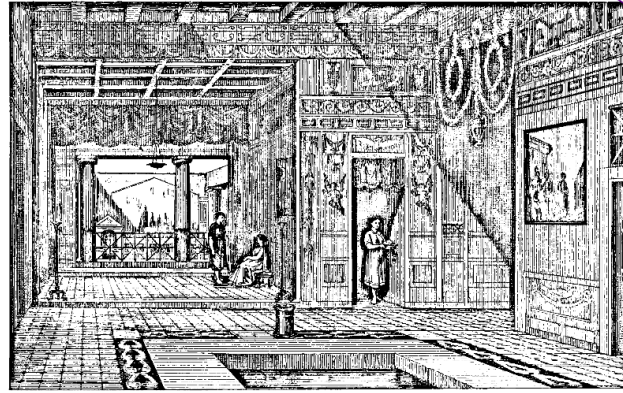


32. Statue vom Dom zu Bamberg.

Zu den reifsten Schöpfungen der deutschen romanischen Plastik gehören die im Westchor des Naumburger Domes um 1270 aufgestellten Statuen von 12 fürstlichen Stiftern und Stifterinnen. Eine von diesen stellt 30 dar. Es ist eine Fürstin in Witwentracht, in einem Buche blätternd. Dies Grundmotiv ist für die Gesamtanlage der Figur entscheidend. Die r. Hand hat, indem sie das Buch hält, zugleich den Mantel aufgenommen, so daß er auf dieser Seite in freien großen Faltenzügen das Buch selbst umschreibt und so das Motiv nach unten ausklingen läßt, in scharfem Kontrast zu den senkrechten Falten ihrer l. Seite, die tief herabfallend sich der bereits gotischen Säule angleichen. Das Antlitz ist durchaus realistisch gehalten, doch ist Porträtmäßigkeit durch die Umstände ausgeschloffen. Idealliert dagegen erscheint das gekrönte Haupt (2 Haken sind abgebrochen) einer Reiterlatz am Dom zu Bamberg, in dem man Konrad III. vermutet (32). Der von einem schlanken Hals getragene, etwas rechts und aufwärts gewendete, von wallendem Haar umrahmte Kopf ist von außerordentlicher Kraft und Frische der Auffassung. — Unruhig wirkt gegenüber der Naumburger Statue das Doppelgrabmal Heinrichs des Löwen und seiner Gemahlin Mathilde im Braunschweiger Dom (31; s. 20). Die Figuren sind, obwohl sie liegen, doch wie stehend gebildet; andererseits aber fallen die Gewänder auch wieder nach dem Geleis der Schwere; so erklärt sich die etwas ängstliche Haltengung besonders des Frauengewandes. Man erkennt hieraus, daß der Künstler gewohnt und geübt war, Standbilder wie etwa 30 zu schaffen; hier sah er sich nun vor eine ganz neue Aufgabe gestellt, die er nicht widerspruchslos zu lösen vermochte. Porträtmäßigkeit ist kaum anzunehmen. Heinrich trägt, wie üblich, in der Rechten das Modell seiner Stiftung, des Braunschweiger Domes.

48

XVIII. Das römisch-griechische Haus.



91

Das altrömische Haus (92) war ursprünglich mit hohem Strohdach versehen; der Rauch des Herdes schwärzte den Mittelraum (daher Atrium) und fand seinen Ausweg durch die Tür und zwei seitliche Flügel, Alae, welche Licht einließen. Das Miteinanderliegen der Grundstücke in städtischen Verhältnissen nötigte sodann, das Dach nach innen zu senken und den Regen durch eine von vier horizontalen Balken getragene Öffnung (Compluvium) in der Mitte des Atriums in ein Bassin (Impluvium) abzuleiten; die für die Lichtzufuhr jetzt entbehrlichen Alae blieben. Dem Eingang (Vestibulum, Fauces) gegenüber befand sich, nach dem Atrium wie nach dem am Ende des Grundstückes liegenden Garten in voller Breite geöffnet, das Geschäftszimmer des Hausherrn, das Tablinum (tabulae). Ein niedriger Oberstod war Regel, zumal für das durch die Alae gebildete Querchiff. Bei größeren Abmessungen wurde das Compluvium durch vier Säulen gestützt; aus dem Atrium tuscanicum wurde dann ein Atrium corinthicum.

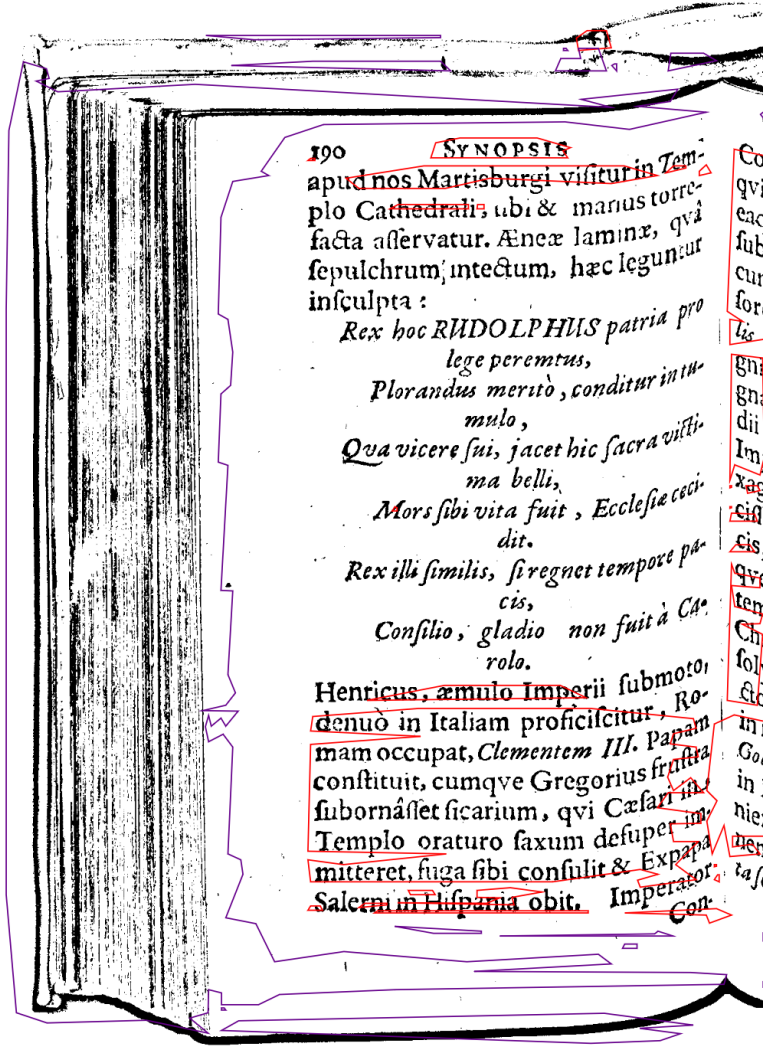
Die griechische Hausanlage hatte sich unterdessen aus einem von dem Megaron beherrschten Hof, welchen offene Hallen und gedeckte Gelasse umgaben, allmählich zu einem regelmäßigen Grundriß entwickelt: dem römischen Atrium entsprach das Peristyl, auf welches alle Räume des Hauses mündeten; das Megaron behielt seine Lage auf der dem Eingang gegenüberliegenden Seite, wurde aber als Besuchszimmer (Exedra), dem Tablinum entsprechend, in ganzer Breite nach dem Peristyl geöffnet und mit den übrigen Räumen unter einem Dache vereinigt. Bei größeren Häusern reisten sich zwei solcher peristylen Anlagen, vorn das Männerhaus, hinten das Frauenhaus, aneinander.



92

Das römisch-griechische Haus, wie es uns Pompeji in den mannigfachen Spielarten aufbewahrt hat, war eine Zusammenordnung der römischen und der griechischen Grundform: an Stelle des Hortus trat das griechische Peristyl. Ein solches Haus zeigt 91. Anwiefern ist hier das Atrium unregelmäßig? Durch das Tablinum blickt man in das Peristyl, welches hier jedoch nicht an der vierten Seite herangeführt ist.

Zeichn. von Adolf Naumann in Leipzig.



190 **SYNOPSIS**
 apud nos Martisburgi visitur in Tem-
 plo Cathedrali, ubi & manus torrefacta
 aservatur. Aeneae laminæ, quæ
 sepulchrum, inter cetera, hæc leguntur
 insculpta :

Rex hoc RUDOLPHUS patria pro
 lege peremptus,
 Plorandus merito, conditur in th-
 mulo,
 Qua vicere sui, jacet hic sacra visiti-
 ma belli,
 Mors sibi vita fuit, Ecclesia cecidi-
 dit.
 Rex illi similis, si regnet tempore pa-
 cis,
 Consilio, gladio non fuit à Cæ-
 rolo.

Henricus, armulo Imperii submoto,
 denuò in Italiam proficiscitur. Ro-
 mam occupat, Clementem III. Papan
 constituit, cumque Gregorius fracta
 subornasset sicarium, qui Cafari in
 Templo oraturo saxum desuper im-
 mitteret, fuga sibi consulit & Expapa-
 Salerni in Hispania obit. Imperator

Co
 qui
 eac
 sub
 cur
 for
 tis
 gn
 gn
 dii
 Im
 xae
 cia
 cis
 qu
 ten
 Ch
 fol
 et
 in
 Go
 in
 nie
 mer
 ta

Fazit und Ausblick

- Lernt gut, generalisiert schlecht
- Ansatz (für dieses Projekt) sehr aussichtsreich
 - Trainingsdaten lassen sich sehr schnell erstellen
 - Material weitestgehend unproblematisch
- Weiter breit Trainingsdaten erstellen
- Gezieltes Active Learning basierend auf Qualitätssicherung (kommt gleich)
- Ggf. Regionentypen schrittweise wieder weiter ausdifferenzieren

Agenda

1. Allgemeine Projektinfos
2. Genereller (OCR) Workflow
3. Layoutanalyse
- 4. OCR-Qualitätssicherung und –Optimierung**
5. NER und NEL

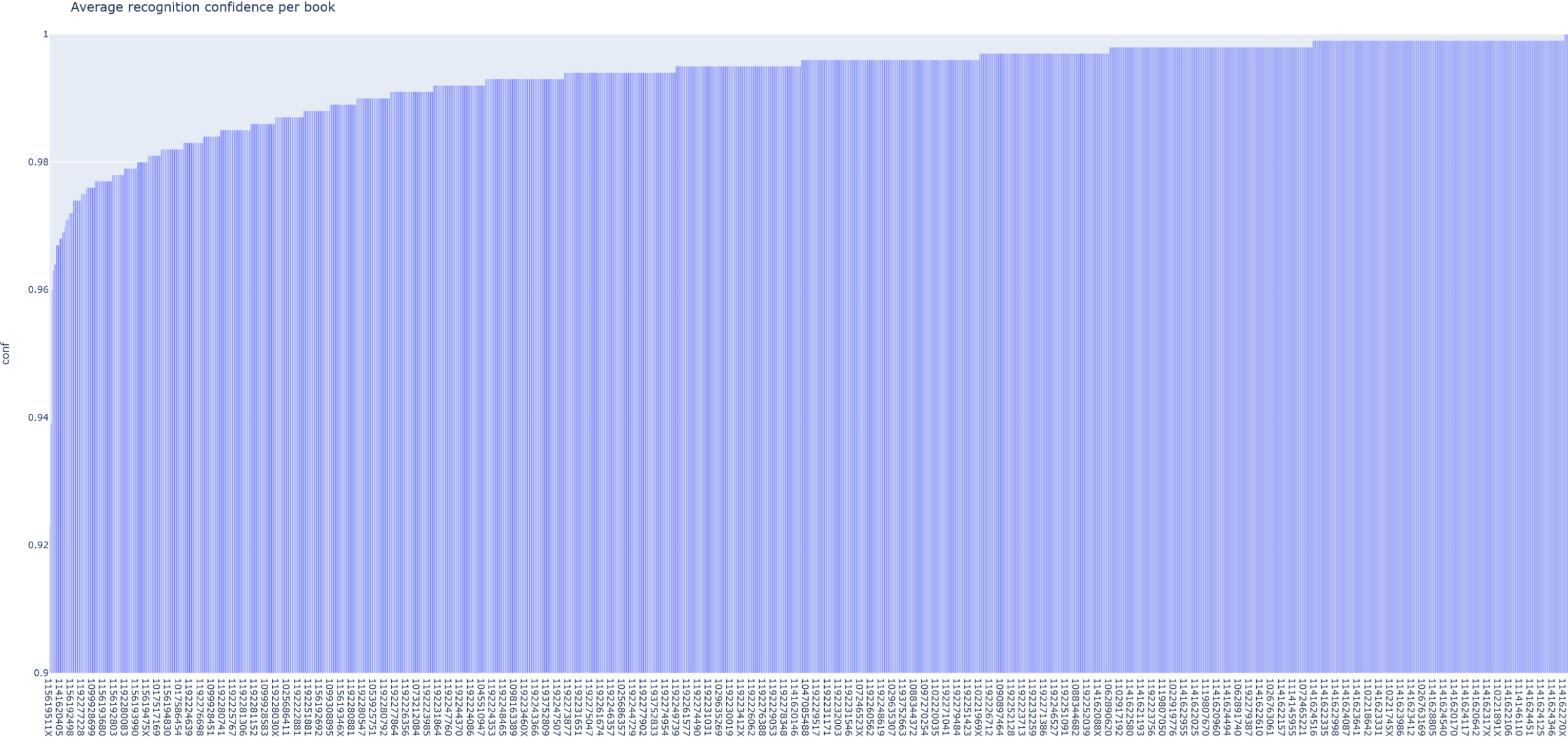
Überblick

- Eingehende manuelle Prüfung aller Werke/Seiten nicht gangbar
- Stattdessen gezielte Prüfung, basierend auf Qualitätsabschätzung
 - Auf Werk-/Seitenebene → jeweils sortiert nach (vermuteter) „Qualität“
 - Flexibel und experimentell!
- Art des Eingriffs dann je nachdem
 - Anpassung von Parametern, Anwendung eines anderen Modells, ...
 - Gezieltes Nachtraining auf Werk-/Clusterebene (Active Learning!)
 - ...?
- Frage: Wie Qualität abschätzen?
 - Bereits gute Erfahrungen mit OCR-Konfidenz, allerdings auch klare Schwächen
 - Nur pragmatische (und idealerweise direkt einsetzbare) Lösungen sinnvoll

Rückblick: Anwendungsfall „Fotidis‘ Schundhefte“

- Material: Deutsche Heftrromane, hauptsächlich 19. Jh.
- Ziel-Anwendung: Quantitative Analyse / Computational Literary Studies
→ (sehr) gute aber nicht perfekte Qualität erforderlich
- Workflow damals
 - Preprocessing und Layoutanalyse mit OCRopus (teils recht fehlerhaft)
 - Texterkennung mit Calamari
 - Erfassung der Calamari-Konfidenz; Aggregation auf Werk- und Seitenebene
 - Prüfung/Korrektur in LAREX
 - Seiten aufsteigend nach Konfidenz sortiert
 - Abbruch, sobald 2-3 Seiten (quasi) korrekt / gut genug
- Hocheffizient!
- Aber: Gerade Layoutanalyse mittlerweile deutlich weiter

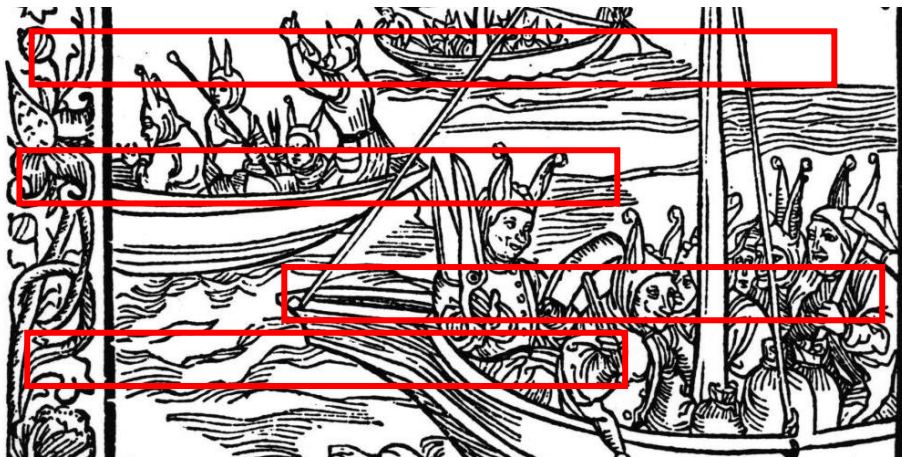
Rückblick: Anwendungsfall „Fotidis‘ Schundhefte“



Stärken/Schwächen OCR-Konfidenz

Text Text Text Text Text

teilsegmentierte Zeile, Zeilen in Bild
→ OCR-Konfidenz klarer Indikator



Lorem ipsum dolor sit gubergren, no sea
amet, consetetur takimata sanctus est
s adipscing elit, sed Lorem ipsum dolor sit
diam nonummy eirmod amet. Lorem ipsum dolor
tempor invidunt ut sit amet, consetetur
labore et dolore magna sadipscing elit, sed diam
aliquyam erat, sed diam nonummy eirmod tempor
voluptua. At vero eos et invidunt ut labore et
accusam et justo duo dolore magna aliquyam
dolores et ea rebum. erat, sed diam
Stet clita kasd

Übersehene Spaltentrennung
→ OCR-Konfidenz ggf. top,
Ergebnis jedoch Quatsch

Fazit und Ausblick

- Gezielte Korrektur aussichtsreich ...
- ... gerade im Zusammenspiel mit Active Learning der Layoutanalyse
- OCR-Konfidenz kann guter Indikator sein ...
- ... allerdings stark abhängig von Art der Fehler

- Einsatz von / Kombination mit NLP-Methoden naheliegend und vielversprechend

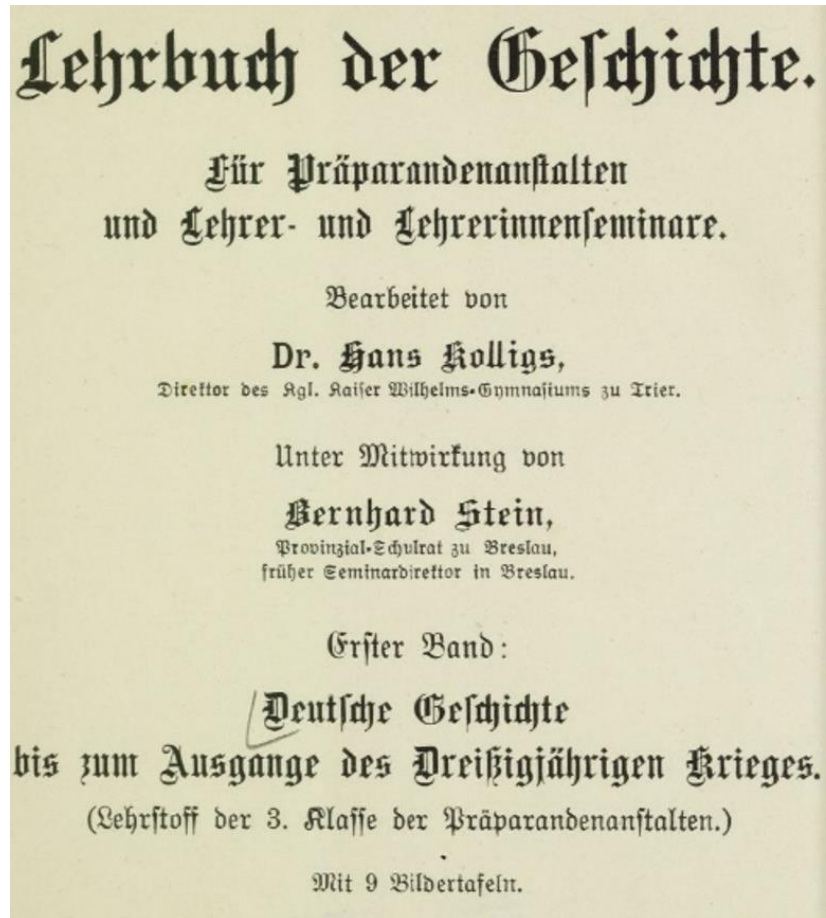
Agenda

1. Allgemeine Projektinfos
2. Genereller (OCR) Workflow
3. Layoutanalyse
4. OCR-Qualitätssicherung und –Optimierung
- 5. NER und NEL**

Überblick

- Systematische Erfassung von Akteuren aus der historischen Schulbuchproduktion unter Einsatz von Named Entity Recognition (NER) und Named Entity Linking (NEL)
- Named Entity Recognition (NER)
 - Benannte Entitäten in natürlichsprachigen Texten automatisiert erkennen und in vordefinierte Kategorien einordnen
 - Standard-Unterscheidung in vier Klassen: Person (PER), Location (LOC), Organisation (ORG), Miscellaneous (MISC); je nach Modell auch andere Klassifizierung möglich
- Named Entity Linking (NEL)
 - Verlinkung/Verknüpfung von Entitäten zu anderen Quellen, z. B. einer Normdatei oder einer anderen zuvor festgelegten Wissensbasis (GND, Wikipedia, ...)
 - Entität eindeutig identifizieren und gegenüber gleichlautenden Entitäten disambiguieren

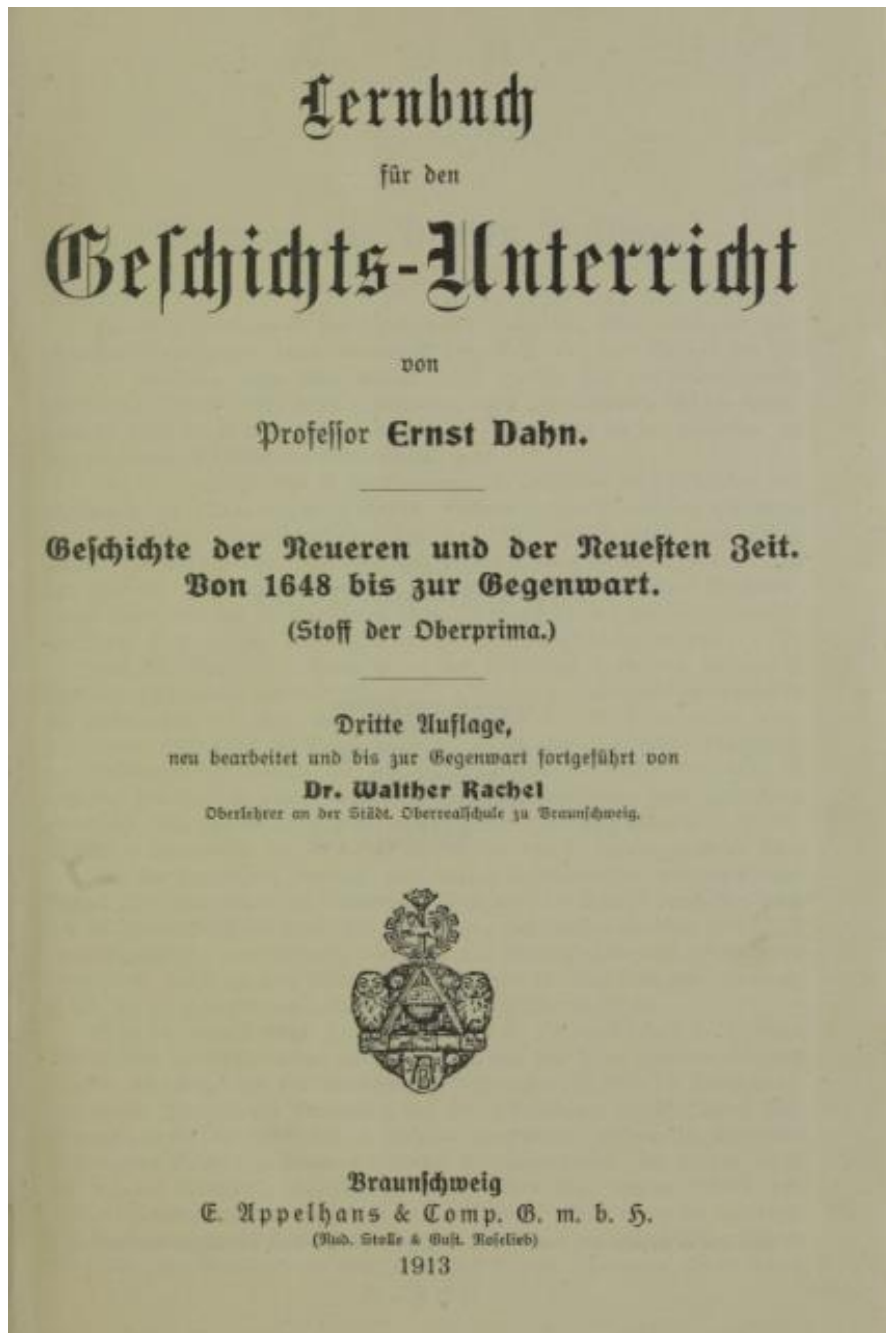
Beispiel NER-Ergebnis auf fehlerhafter OCR



Lehrbuch der Geschichte MISC .
Für Präparandenanstalten
und Lehrer- und Lehrerinnenseminare.
Bearbeitet von
Dr. Hans Kolligs PER .
Direktor des Kal ORG . Kaiser Wilhelms-Gymnasiums ORG zu Trier LOC .
Unter Mitwirkung von
Vernhard Stein PER .
Provinzial. Schulrat zu Breslau LOC ,
früher Seminarlehrer in Breslau LOC .
Erster Band:
Deutsche Geschichte
bis zum Ausgange des Dreißigjährigen Krieges MISC .
(Lehrstoff der 3. Klasse der Präparandenanstalten.)
Mit 9 Bildertafeln.

Workflow

- Plain Text erfassen
 - Mittels OCR erzeugen
 - Aus bestehenden Ergebnissen extrahieren (für Teilkorpus liegen alte ABBYY Ergebnisse vor)
- Fokus auf Textabschnitte, die besonders viele relevante Entitäten enthalten
 - Titelblätter, Vorworte, ...
 - Gezielte Durchführung, basierend auf Infos aus METS
- NER: Modelle identifizieren Entitäten
- NEL: Unterstützte Disambiguierung und Verknüpfung mit externen Quellen (GND)
- Verknüpfung mit den Ursprungsdaten (ALTO)
- (Visualisierung, ggf. Netzwerkanalyse etc.)



```
00000005.xml X
PcGts Page ReadingOrder
1 <?xml version='1.0' encoding='utf-8'?>
2 <PcGts xmlns="http://schema.primaresearch.org/PAGE/gts/pagecontent/2019-07-15"
2 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
2 xsi:schemaLocation="http://schema.primaresearch.org/PAGE/gts/pagecontent/2019-07-15
2 http://schema.primaresearch.org/PAGE/gts/pagecontent/2019-07-15/pagecontent.xsd">
3 <Metadata>
4 <Creator>xml_converter</Creator>
5 <Created>2024-10-01T10:09:55.392561</Created>
6 <LastChange>2024-10-01T10:09:55.395534</LastChange>
7 </Metadata>
8 <Page imageFilename="00000005.png" imageWidth="2254" imageHeight="3618">
9 <ReadingOrder>
10 <OrderedGroup id="g0">
11 <RegionRefIndexed index="0" regionRef="r0"/>
12 <RegionRefIndexed index="1" regionRef="r1"/>
13 <RegionRefIndexed index="2" regionRef="r2"/>
14 <RegionRefIndexed index="3" regionRef="r3"/>
15 <RegionRefIndexed index="4" regionRef="r4"/>
16 <RegionRefIndexed index="5" regionRef="r5"/>
17 <RegionRefIndexed index="6" regionRef="r6"/>
18 <RegionRefIndexed index="7" regionRef="r7"/>
19 </OrderedGroup>
20 </ReadingOrder>
21 <TextRegion id="r0" type="paragraph">
22 <Coords points="696,338 1246,348 1246,508 1074,562 1074,619 1858,692 1858,908
22 1026,1012 1026,1043 1372,1155 1362,1237 586,1227 586,1155 922,1043 922,1012 94,908
22 94,692 872,619 872,562 686,508 686,348"/>
23 <TextLine id="r0_l0">
24 <Coords points="696,352 772,352 784,388 836,388 844,384 904,384 912,384 980,384
24 996,352 1048,352 1060,384 1132,384 1144,348 1236,348 1144,348 1236,348 1236,508
24 1144,508 1236,508 1144,508 1132,472 1060,472 1048,476 996,476 980,476 912,476 904,476
24 844,476 836,476 784,476 772,508 696,508 696,352"/>
25 <Baseline points="696,472 1236,472"/>
26 <TextEquiv>
27 <Unicode>Kernbuch</Unicode>
28 </TextEquiv>
29 </TextLine>
30 <TextLine id="r0_l1">
31 <Coords points="882,562 900,562 902,564 930,564 933,576 953,576 954,562 985,562
31 986,562 1010,562 1016,575 1034,575 1038,575 1064,575 1064,606 1038,606 1034,607
31 1016,607 1010,607 986,607 985,607 954,607 953,607 933,607 930,607 902,607 900,619
31 882,619 882,562"/>
32 <Baseline points="882,607 1064,607"/>
33 <TextEquiv>
34 <Unicode>für den</Unicode>
35 </TextEquiv>
```

Kernbuch
für den
Geschichts-Unterricht
von
Professor Ernst Dahn.
Geschichte der Neueren und der Neuesten Zeit. Von 1648 bis zur Gegenwart.
(Stoff der Oberprima.)
Dritte Auflage,
neu bearbeitet und bis zur Gegenwart fortgeführt von Dr. ttialtber Kachel
Oberlehrer an der Stöbt. Oberrealschule zu Braunschroieg.
Braunschweig
E. Appelhans & Comp. 05. m. b. H.
(Rud. Stolle & Gust. Roselieb)
1913

Kernbuch
für den
Geschichts-Unterricht
von
Professor Ernst Dahn **PER gnd** .
Geschichte der Neueren und der Neuesten Zeit. Von 1648 bis zur Gegenwart.
(Stoff der Oberprima.)
Dritte Auflage,
neu bearbeitet und bis zur Gegenwart fortgeführt von **Dr PER** , **ttialtber Kachel PER**
Oberlehrer an der Stöbt. **Oberrealschule ORG** zu **Braunschroieg LOC gnd** .
Braunschweig E. Appelhans & Comp ORG . 05. m. b. H.
(**Rud PER** . **Stolle PER** & **Gust PER** . **Roselieb PER**)
1913

Erste kleine Experimente

- Aktuell basierend auf bereits vorliegenden ABBYY Ergebnissen, mit eher überschaubarer Qualität
- Testen verschiedener Standard-Modelle
 - Spacy https://spacy.io/models/de#de_core_news_lg
 - Flair <https://github.com/flairNLP/flair>
 - Babelscape <https://huggingface.co/Babelscape/wikineural-multilingual-ner>
- Erste kleine Auswertung
 - Menschlich erzeugte Golddaten mit Fokus auf Personen und Orte (173 Entitäten aus drei Werken)
 - Datengrundlage sind die Scans, nicht das erzeugte OCR
 - Flair Modell liefert beste Ergebnisse (Precision >0,9; Recall >0,8)
 - Exakte Evaluation schwierig → grobe Werte als erste Abschätzung!

Fazit und Ausblick

- Noch einige Herausforderungen, u. a.
 - Teils sehr fehlerhafte OCR
 - Umgang mit langen zusammengesetzten Entitäten
„Frau Akademische Oberrätin Priv. Doz. Dr. habil. phil. Katrin Dennerlein (geb. Fischer)“
 - Disambiguierung (Herr Müller ≠ Herr Müller ≠ Herr Müller)
- Prinzipiell aber erstaunlich robust, performant und aussichtsreich
- Verschiedene Möglichkeiten, Ergebnisse erheblich zu verbessern, u. a.
 - Anwendung auf bessere OCR
 - Bessere NER Modelle, ggf. durch Finetuning / Domain Adaptation
 - Ergänzung durch heuristische Methoden
 - NEL erheblich unterstützten (GND-API + eigene GUI)

Vielen Dank für die Aufmerksamkeit!

Projektbeteiligte:

- Fabian Dombrowski (NER und NEL, GEI)
- Janik Haitz (Layoutanalyse und Texterkennung, ZPD)
- Corinna Keupp (NER und NEL, ZPD)
- Florian Langhanki (Layoutanalyse und Texterkennung, ZPD)
- Johannes Leitgeb (NER und NEL, ZPD)
- Sebastian Klaes (Koordination, GEI)
- Anke Hertling (Antragstellerin, GEI)
- Christian Reul (Antragsteller, ZPD)