

# Use Case *Anagnosis*

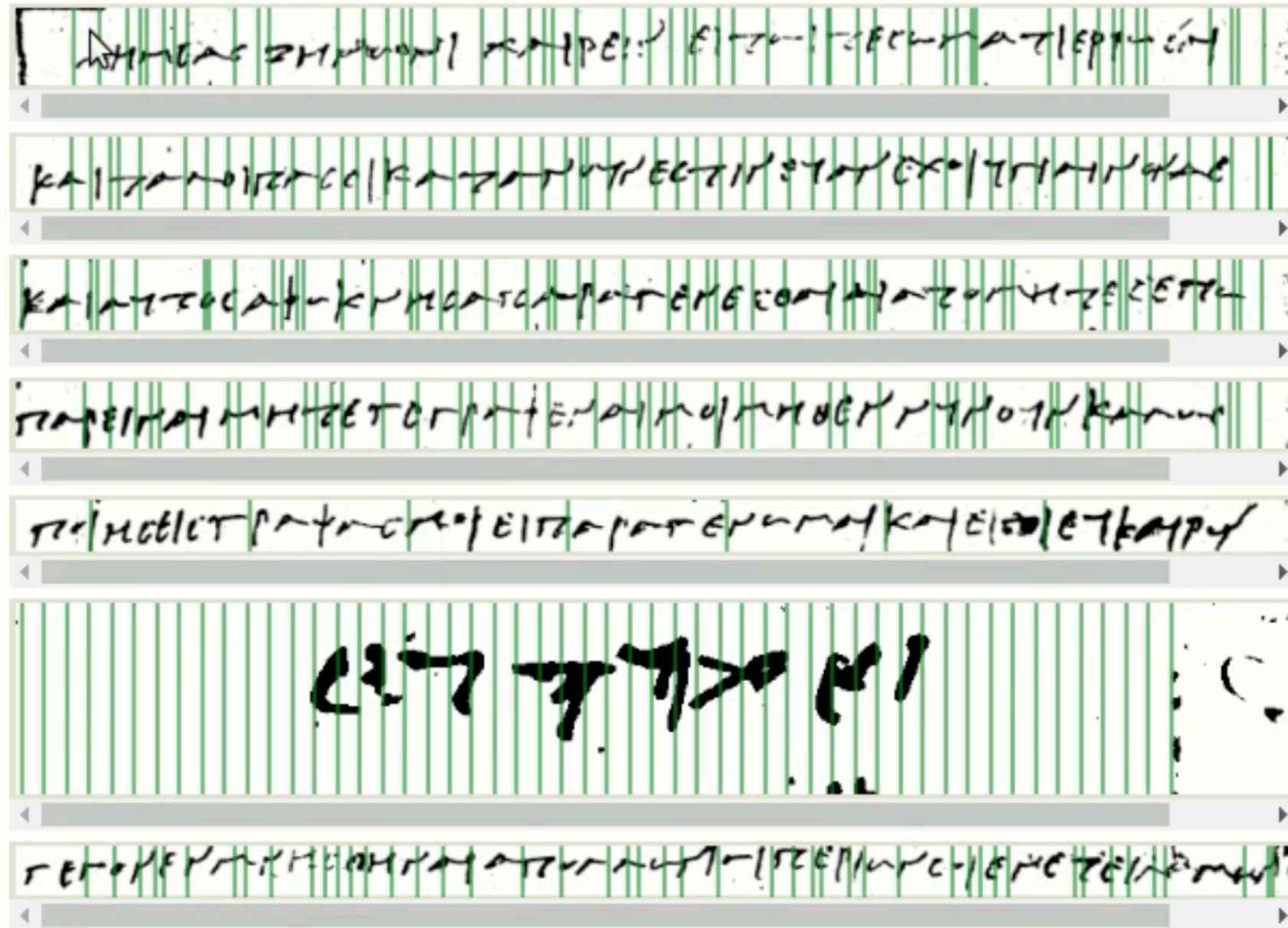
OCR-Optimierung und *Sequence Alignment* bei griechischen Drucken  
aus der Offizin des Aldus Manutius (1449-1515)

Prof. M. Erler, Dr. H. Essler, Vincenzo Damiani, M.A.  
Institut für klassische Philologie, Lehrstuhl I (Gräzistik)



Please use the handles (Blue Lines) on images to set the alignments. When Done Press "Save" to perform the operation

Delete Save



Δμίας Ζήνωνι χαίρειν. εἰ τῷ τε σώματι ἔρρωσαι

καὶ τὰ λοιπὰ σοι κατὰ νοῦν ἐστίν, εὖ ἂν ἔχοι· ὑγίαινον δὲ

καὶ αὐτός. ἀφώκνησα(\*) παραγενέσθαι διὰ τὸ μήτε σέ πω

παρεῖναι μήτε γεγραφέναι μοι μηθέν. νῦν οὖν καλῶς

ἔποιήσεις γράψας μοι εἰ παραγένωμαι καὶ εἴ σοι εὐκαιρον

γέγονεν μνησθῆναι Ἀπολλωνίωι περὶ ὧν σοι ἐνετειλάμην.

εὐτύχει.

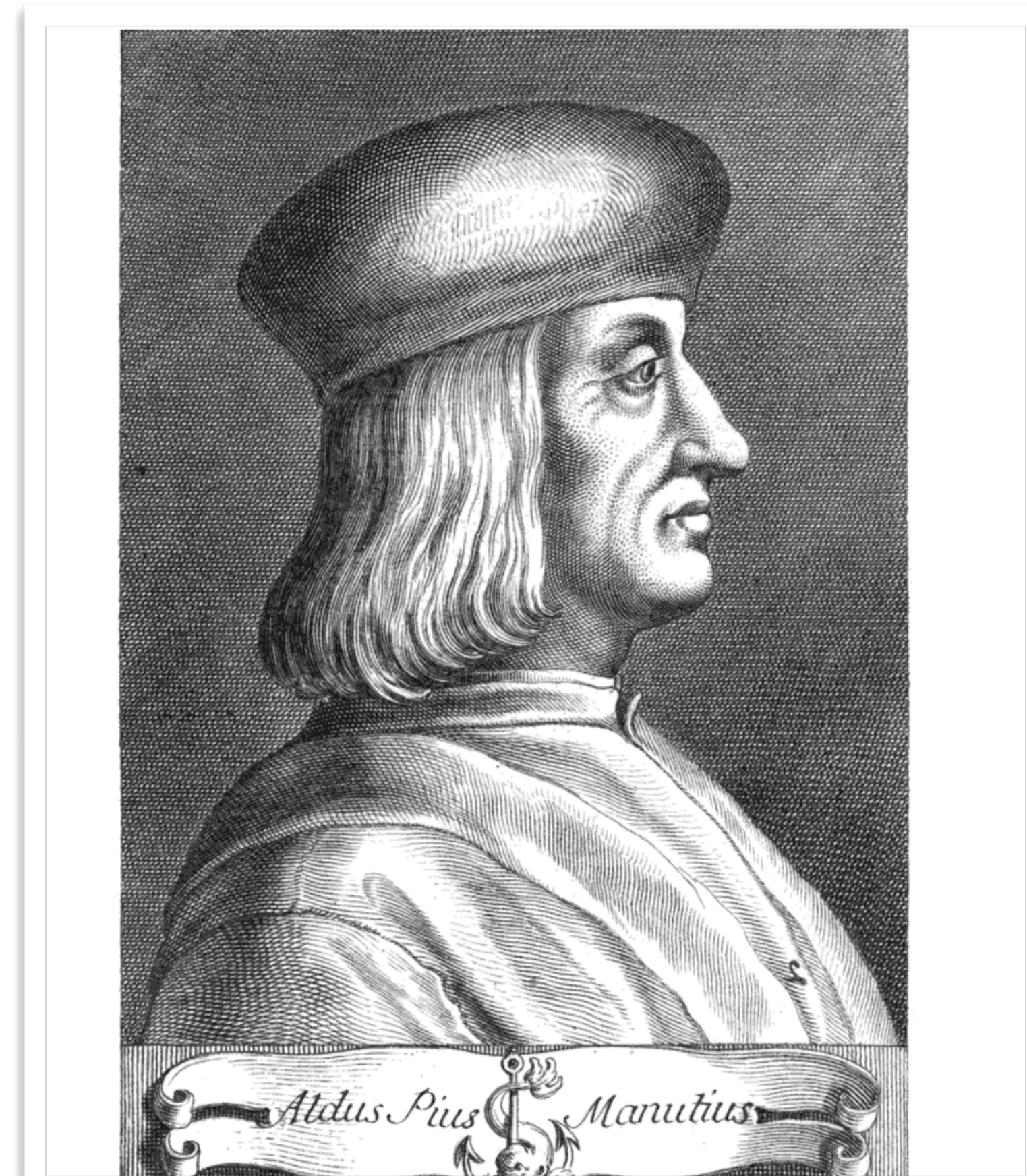
© This Tool is constructed by DFKI GmbH. See Contact Details.

# Anagnosis 1.0: Bild-Text-Alignment bei griechischen Papyri


## Modul Synoptischer Editor

# Aldus Manutius (1449-1515)

- Verleger in Venedig (1495 -1515)
- *Editiones principes* griechischer Klassiker (Aristoteles [1495-98], Aristophanes [1498])
- Wichtige Neuerungen im Druckwesen
- Seine Editionen dienten als Vorlage für spätere Druckausgaben



Aristoteles, *Analytica priora*  
(1495-1498)  
ISTC ia00959000; GW 2334



ἈΡΙΣΤΟΤΕΛΟΥΣ ἈΝΑΛΥΤΙΚῶΝ ΠΡΩΤΕΡῶΝ  
ΠΡΩΤΟΝ.  
ΠΕΡΙ ΤῶΝ ΤΡΙῶΝ ΣΧΗΜΑΤῶΝ.

**Π**ΡΩΤΟΝ εἶπεν περὶ τίνος ἢ σκέψις ὅτιν ὅτι περὶ ἀποδείξειν, καὶ ἐπισημῆς ἀποδεικτικῆς· εἶτα διορίσαι, τί ἐστὶ πρότασις, καὶ τὸ ὅρος, καὶ τὴν συλλογισμόν· ἢ ποῖος τέλειος ἢ ποῖος ἀτελής· μετὰ δὲ ταῦτα, τί τὸ ἐν ὅλῳ εἶναι, ἢ μὴ εἶναι, τὸ δὲ τὰς δευτέρας λέγομεν τὸ κατὰ παντός ἢ μηδενός κατηγορεῖσθαι. Πρότασις μὲν ἔνδεξις λόγος καταφατικός ἢ ἀποφατικός, τίνος κατατίνοσ· οὐ γὰρ ἢ καθόλου, ἢ ἐν μέρει, ἢ ἀδιόριστος· λέγω δὲ καθόλου μὲν τὸ, παντὶ ἢ μηδενὶ ὑπάρχειν, ἢ μὴ ὑπάρχειν, ἀδιόριστον δὲ, τὸ ὑπάρχειν ἢ μὴ ὑπάρχειν, ἀναλόγως καθόλου ἢ κατὰ μέρος· οἷον τὸ ἐναντίων, εἶναι αὐτὸ ἐπισημίω· ἢ τὸ πῶν ἢ δυνάμει, μὴ εἶναι ἀγαθόν· διαφέρει δὲ ἢ ἀποδεικτικὴ πρότασις, τῆς διαλεκτικῆς, ὅτι ἢ μὲν ἀποδεικτικὴ, λήψις θατέρου μορίου τῆς ἀντιφάσεως ὅτιν τὸ ἄρρωστῶν, ἀλλὰ λαμβάνει ἀποδεικνύων· ἢ δὲ διαλεκτικὴ, ἐρώτησις τῆς ἀντιφάσεως ὅτιν· οὐδὲν δὲ διοίσει πρὸς τὸ γινώσθαι τὸ ἕκαστον συλλογισμόν· καὶ γὰρ ὁ ἀποδεικνύων καὶ ὁ ἐρωτῶν, συλλογίζεται, λαβὼν τι κατατίνοσ ὑπάρχειν, ἢ μὴ ὑπάρχειν· ὥστε ἐστὶ συλλογιστικὴ μὲν πρότασις, ἀπὸ τῶν καταφάσεως ἢ ἀποφάσεως τίνος κατατίνοσ, κατὰ τὸν εἰρημέρον τρόπον· ἀποδεικτικὴ δὲ ἐὰν ἀληθὴς ἢ, εἰ δὲ ἀτὸν ἀρχῆς ὑποθέσεων εἰλημμένη· δια-

ANGELI POLITIANI EPISTOLARVM LIB. PRIMVS.

Angelus Politianus Petro Medici suo. S. D.

GISTI MECVM Sæpenumero Magnani-  
me Petre Medices, ut colligerem meas episto-  
las, & in uolumen redactas publicarem. Colle-  
gi, ne non in oibus obsequer tibi, quo sunt  
in uno, spes omnes, o pesq; meæ sitæ, Neq; col-  
legi tamen uniuersas, id enim laboriosius, q̄ Si-  
byllæ folia. Non scripseram uidelicet ad hoc,  
ut i unum corpus referretur, sed ad usum præ-  
sentem duntaxat, oblati argumentis, nõ quæ-  
sitis. Ita nec exempla mihi retinui, nisi quarundam, puto minus feli-  
cium, quæ diu iam cū blattis & tineis rixabatur. Verūtamen, ut instar  
uoluminis efficerem, nõnullas etiam missas ad me, sed a doctis dunta-  
xat, inserui, quæ marcentē quasi stomachum lectoris excitarent. Est au-  
tem omnino stilus epistolæ mearum, ipse sibi dispar, Quo noie mul-  
tum quoq; scito reprehendar. Nam nec eadem mihi semper uoluntas,  
nec idem cuiq; uel personæ, uel materiæ cõgruebat. Non deerunt ergo  
qui dicant, ubi tam uarias epistolas legerint ( si qui modo legerint ) ite-  
rum me Miscellanea, non epistolas composuisse. Sed inter diuersas opi-  
niones, & præcipientiũ de epistolis, & epistolas scribentiũ, speraui fore  
perfecto, nec ubi mihi patrociniũ deesset. Occurreret aliquis forsan, qui  
Ciceronianas esse neget, huic ego dicã ( nec sine auctore tamen ) in epi-  
stolari stilo silendũ prorsus esse de Cicerone. Rursus alius hoc ipsum  
culpabit, quod æmuler Ciceronẽ, sed respondebo, nihil mihi esse ma-  
gis in uotis, q̄ ut uel umbram Ciceronis assequar. Optaret alius, ut ora-  
torem Plinium sapere, quod huius & maturitas, & disciplina laudatur.  
Ego cõtra totũ illud aspernari me dicã Plinii sæculum, Sed & si Pliniũ  
cuiq; redolebo, tuebor ita me, quod Sidonius Apollinaris, nõ omnino  
pessimus auctor, palmã Plinio tribuit in epistolis. Symmachum si cui  
referre uidebor, nõ pudebit, ut cuius & breuitas celebretur, & rotundi-  
tas. A besse rursus a Symmacho si cui credar, negabo mihi siccitatẽ pla-  
cere. Longiores quædam dicuntur epistolæ, tales Plato scripsit, Aristote-  
les, Tucydides, Cicero. Dicuntur aliæ contra nimis breues, obiciã tunc  
ego Dionem, Brutum, Apollonium, Marcũ, Philostratum, Alciphro-  
nem, Iulianum, Lybanium, Symmachum, sed & Lucianum, quẽ falso  
Phalarim uulgo putant. Damnabit alius me, quod argumẽta non sint

Politianus, *Epistulae* (1498)  
ISTC ip00886000; GW M34727

# Material

Drei griechische Schrifttypen

(Bezeichnung nach TW)

Type 1:146Gr

Θ. Λῆς πρὶ τῶν νυμφῶν λῆς αἰπρὲς τῆδε καθόζας  
ὧς τὸ κάταρτος τοῦτο γέλωφορᾶτε μυεῖκας,  
Συείσθεν, τὰς δ' αἰγὰς ἐγὼν ἐν τῷδε νομοσῶ;  
Αἱ. Οὐ θέμις ὧ πρὶ μὲν τὸ μεσαμβρινόν, ὄθεμις ἄ μὲν  
Συείσθεν, τὸν πᾶνα δευδίκασος. ἢ γὰρ αἰπύρας

Ωμίλη, πθίουσα παλαῖων ἔθια λαῶν.  
Αλλ' ἔμπερς ἐπὶ κεινω κατὰ ρύριον γγύοσ ἔερρ  
Ἡρχετο δ' ὄξ ὀρίων ἐπιδείλιος ἠχασήτωι  
Μουάξ. οὐδὲ τῷ ἐπεμίσητο μφλιχίοισι  
Αλλ' ὅπ' τ' ἀνθρώπων μεγάλασ πλήσαιτο κελῶνας

Type 7:114Gr

Ἰ μὲν, ἐπιγράφουσι περὶ θανα  
σίμων φαρμάκων. οἱ δὲ, ἀντι  
φάρμακον. ἄλλοι ὃ, ἀλεξιφάρ  
μακον, ὡς κἢ αὐτὸς φησι, Ρ εἶα κέ  
τοι ποσίεσιν ἀλέξια φαρμακο

Type 9:84Gr

# Vorgesehene Arbeitseinheiten

1. OCR-Optimierung für griechische Zeichen bei bestehenden Digitalisaten
2. Erstellung neuer Digitalisate
3. *Sequence Alignment* mit bestehenden Textdatenbanken

gtc.informatik.uni-wuerzburg.de

μέλημα, πρὸς διόρθωσιν γὰρ ἀγαγὼν αὐτὸν αἰτεῶς γενήσῃν τῷ μὴ

γεγονέναι καθάπερ ἐγὼ τοῦ γεγονέναι. Νουθετικὸς δὲ ἔστιν ὅς καὶ δια

ον ἀξίωσον τὸ πρᾶγμα. καὶ γὰρ σὸ πάλιν ὑπᾶλλου παπθῶν,ρα-

δύς ἀν νεγκας. τὸ δὲ ἔσαντὸν δικαίον· μὴ δόκειον μήτε γον

βι μὲν ὑπολαμβάνεις μηδεμίαν ὑφέξεν δίκην ἐφοῖς διαπέπρα

ξιαι, πρᾶττε. σθι δὲ ἀκριβῶς ὡς οὔτε ἀναπτησόμενο, οὔτε κατα

ιδεῖναπὼς κέχρηται τοῖς ἐγκεχειρισμῶοις ἀνελευθέρως καὶ

τῆς προαιρέσεως ἀναξίως, καὶ ἐγὼ σιωπῶ παρ ἄλλων ἀκούση· γρά

τίαι τῶν πορρόρητων εἰπεῖν, ὡς ἀνόμοιόν ἐστιν, ἐν παχεῖ σώματι

Virtual Keyboard

Image-height: 28  Adjust textheight  
 Text-height: 28  
 Vk-height: 10   
 FC Font-size: 10

Shown lines: 30

# Weboberfläche zur OCR-Optimierung



# *Sequence Alignment*


- Durch OCR erkannter Textabschnitt aus der *Aldina*
  - Volltextdatenbank
- 



---

Parallelanzeige: erkannter Textabschnitt  
verglichen mit möglichen  
Entsprechungen aus der Datenbank

Automatische  
Erstellung von  
Ground Truth

A large, curved arrow on the left side of the diagram, pointing from the bottom text block up to the top list of items.

# OCR und *Sequence Alignment*: Anwendungsbereiche

- Übertragung auf Handschriften
- Automatisierte Vorarbeiten zur Kollation
- Identifizierung handschriftlicher Vorlagen für frühneuzeitliche Druckausgaben
- Durchsuchbarkeit und Typenklassifizierung

ὡς κ  
τοιο.

διὰ βίβριου ἔσ. αἰδὲ γὰρ ἡ μάρισα πᾶσι πράξει  
ἢ θ' εὐρήσῃ τὰ κατὰ ἄετην, ἢ τὰς τύχας οἷσ' ἀμύσει  
ἢ πάντῃ. πᾶσι ἐμειλώσῃ. οὐ γὰρ ὡς ἀγαθός, ἢ ἕξαι  
ζῶνος ἀνδρὸς φόρου. πρῶτον δὲ γίνονται κατὰ τύχην  
ἢ διαφροσύντων μεγέθει καὶ μικρότητι, τὰ μὲν μί  
κράτων εὐτύχημάτων, ὁμοίως δὲ ἢ τῶν ἀντιθέτων,  
δηλον ὡς οὐ ποιεῖ εὐπὴν τῆς ζωῆς. τὰ δὲ μεγάλα  
ἢ πρῶτον γίνονται.

Immanuel Rhusotas' Handschrift  
Aristoteles, *Ethica Nicomachea*

**Ι**ΠΕΘεὰ κρυφαίων ἐπιμαρτυρα  
λύχρον ἐρώτων.  
Καὶ νύχρον πλωτῆρ θαλασσοτῆ  
ρων ὑμεναίων.  
Καὶ γέμον ἀχνοένταρον, οὐκ ἴδεν  
ἀφθότος ἦως.  
Καὶ σῆσον καὶ ἀβυσσον ὅπῃ γέμοσ' ἐν νυχτὶ Ἡρώϊσ.  
Νηχόμενον τε Λέανδρον ὁμοῦ, ἢ λύχρον ἀκούω

Aldus Manutius  
Musaeus, *Hero et Leander*

# OCR und *Sequence Alignment*: Anwendungsbereiche

- Vergleich zwischen verschiedenen **Druckversionen** desselben Textes
- **Erstzuordnung** noch nicht erforschter Handschriften bzw. Druckausgaben
- Erforschung von **Rezeptionsgeschichte** und **intertextuellen Verhältnissen**

# Aktueller Stand

- Erstellung von Ground Truth und Training eines OCR-Modells
- Erste Versuche bei Anwendung einer *Sequence Alignment*-Software (*Passim*)