

Die Möglichkeiten der automatischen Textergänzung auf Papyri

Michaela Rücker

Abstract

This paper provides an overview of the current work status of the eAQUA subproject for Papyrology. The text divides into three sections, corresponding to the originally target of the subproject: the classification of papyri, the textcompletion and the identification of unknown literary fragments. Presented in more detail is particularly the tool to the textcompletion. The user will be made familiar with the various procedures and the different ways in terms of search strategies. Text-Mining can be of great help here, since it puts together the pieces and finds parallel texts. Many documentary papyri and inscriptions share certain formulas, structures of the text, and keywords, which reveal them as typical public records. Therefore, on the basis of known parallels, they can be completed.

Keywords

Klassifikation – Textergänzung – Verfahren – Kontextsuche – Kookkurrenzen

Papyrus zählt zu den gebräuchlichsten Schreibmaterialien in vielen Kulturen der antiken Mittelmeerwelt. Er wurde u.a. in Ägypten hergestellt und in großen Mengen exportiert. Auch andere Beschreibstoffe wie die Scherben von zerbrochenen Tongefäßen (Ostraka) benutzte man in größerem Maße als preiswerten Ersatz für Papyrus zur Niederschrift kurzer Texte des täglichen Lebens. Das Fach Papyrologie beschäftigt sich mit der Entzifferung, Übersetzung und Interpretation der auf Papyrus und anderen Beschreibstoffen erhaltenen literarischen Texte und dokumentarischen Urkunden in griechischer (seltener in lateinischer) Sprache vorwiegend aus Ägypten. Zum Forschungsgebiet gehören die Papyri von der Eroberung Ägyptens durch Alexander den Großen (332 v. Chr.) bis zur arabischen Eroberung (641 n. Chr.).¹ Die Papyrologie unterscheidet zwischen dokumentarischen und literarischen Papyri. Zu den dokumentarischen, die den weitaus größten Teil der erhaltenen Papyri ausmachen, gehören u.a. Urkunden, Briefe, Verträge und Listen.

Die Arbeit des Papyrologen beinhaltet im ersten und wichtigsten Schritt die Entzifferung des Textes und die Anfertigung einer Transkription – das kann wegen der z.T. extrem kursiven Schrift vieler dokumentarischer Papyri zu einer langwierigen Tätigkeit werden. Mit Hilfe von Paralleltexten wird versucht die häufig im Text vorhandenen Lücken bzw. fehlenden Partien zu ergänzen.

Das Teilprojekt 4.6. „Papyrologie“ verfolgt im Rahmen von eAQUA drei Ziele, die die Aufgabe des Papyrologen – durch Editionen das in den Papyri enthaltene Quellenmaterial aufzubereiten und zur weiteren Auswertung zu erschließen –, erleichtern sollen.

Diese drei Ziele sind:

- eine Klassifizierung der Papyri,
- Textergänzung/Textvervollständigung,
- Identifikation/Eingrenzung unbekannter literarischer Papyri, der so genannten „Wartetexte“², mit Hilfe des Gesamtcorpus der antiken Texte: Auf diese Art und Weise soll die unbekannte Textpassage einem antiken Autor zugewiesen oder zumindest

¹ In diesem Zeitraum war die Verwaltungssprache in Ägypten griechisch.

² Das sind Texte, deren Identifizierung bisher noch nicht gelungen ist.

näher eingegrenzt oder – da die meisten Funde fragmentarisch sind – zumindest ergänzt werden.

Klassifikation

Das erste Ziel, aber auch der erste Schritt für die weitere Arbeit, war die Klassifizierung der Texte nach bestimmten inhaltlichen und formalen Kriterien. Dafür wurde die Klassifikation des Papyrusprojektes Halle – Jena – Leipzig auf die ca. 60.000 Papyri der Duke Databank of Documentary Papyri angewandt. Übernommen wurde diese Klassifikation von Friedrich Preisigke, Sammelbuch griechischer Urkunden aus Aegypten, begonnen 1915, der damit einen internationalen Standard für zerstreut publizierte Papyri geschaffen hat.

Als Trainingsmenge für die automatische Klassifikation dienten die bereits publizierten Papyri und Ostraka aus Halle, Jena und Leipzig. Im nächsten Schritt erfolgte die Evaluierung, indem die signifikanten Terme auf ihre Relevanz hin überprüft wurden. Diese Überprüfung war wichtig, da innerhalb dieser signifikanten Terme die häufigsten und seltensten Wörter in den einzelnen Kategorien für die Gewichtung eliminiert wurden. Aufgrund des Formularcharakters der meisten dokumentarischen Papyri können relevante Wörter durchaus eliminiert worden sein, doch sind häufige Terme wichtig zur Inhaltsbestimmung.

Die Arbeit an der Klassifikation ist innerhalb des Teilprojektes am weitesten fortgeschritten und in einem Stammbaum visualisiert worden, in dem die einzelnen Kategorien zusammen mit den zugeordneten Papyri erkennbar sind.

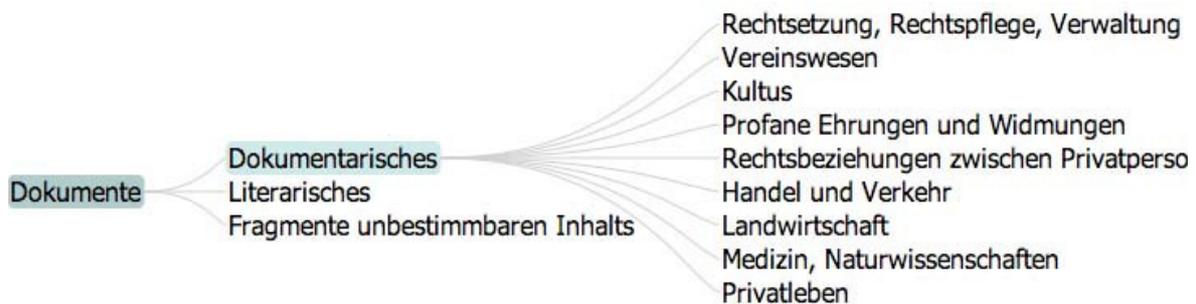


Abb. 1: Klassifikationsstammbaum dokumentarischer Papyri - Ausschnitt I



Abb. 2: Klassifikationsstammbaum dokumentarischer Papyri - Ausschnitt II

Im nächsten Schritt soll die abgeschlossene Klassifizierung in das Textergänzungstool eingearbeitet werden, um die Ergänzungsvorschläge eingrenzen zu können – so weit aus dem fragmentarischen Text hervor geht, welcher Klassifizierungskategorie er angehört.

Textergänzung/Textvervollständigung

Wie in der Einleitung erwähnt, spielt die Textergänzung bei der Arbeit des Papyrologen eine entscheidende Rolle, da die Zeugnisse auf Papyrus in der Regel nie vollständig, sondern als unterschiedlich große Fragmente gefunden werden. Oft kommt es sogar vor, dass Teile ein und desselben Stückes sich in verschiedenen Sammlungen befinden, ohne dass die Zugehörigkeit zueinander erkannt wird. Text-Mining kann hilfreich sein, um Texte wieder zusammenzufügen oder zumindest zu ergänzen, indem sich Paralleltexte finden lassen. Da ein Großteil der dokumentarischen Papyri als formale Urkunden diverser Art jeweils eine ähnliche Textstruktur besitzt, geben sie sich durch bestimmte Schlüsselworte/Keywords als ganz bestimmte Urkundenform zu erkennen. Beispielsweise: ἐμίσθωσεν, ἐκδωκεν/ἔλαβεν, ὁμολογῶ – μισθώσασθαι = Miet/Pachtvertrag; ἀπέδοτο – ἐπρίατο oder ὁμολογῶ πεπρακέναι = Kaufvertrag; ἔδοξεν – δεδόχθαι = Volksbeschluss. Durch die Kenntnis entsprechender Parallelen ist eine Ergänzung möglich.

Das Ergänzungstool liegt in der ersten Version vor, die Textergänzung funktioniert momentan auf ein Wort. Die Testphase des Tools läuft seit Januar 2010, getestet werden vorsätzlich fragmentierte Papyri.

Die Erläuterungen zur Funktionsweise des Ergänzungstools erfolgen am Beispiel eines Leipziger Kaufvertrages (P.Lips. I 2):

ἀπέδοτο Τιτῶς Πατοῦτος Περσίνη ὡς (ἐτῶν) ξ μέση μελίχρω(ς) μακροπρόσω(πος) εὐθύριν μετὰ κυρίου τοῦ ἑαυτῆς πρεσβυτέρου υἱοῦ Ψεννήσιος τοῦ Ἄρσιήσιος Πέρσου τῆς ἐπιγονῆς ὡς (ἐτῶν) ν μέσου μελάνχρω(τος) ὑποκλά(στου) ἀναφαλάντου μακροπρ(οσώπου) εὐθυρίνου, ὀφθαλμὸν δεξιὸν βεβλαμμένου ἀπὸ τῆς ὑπαρχούσης αὐτῆ καὶ τοῖς ἀδελφοῖς γῆς σιτοφόρου ἀδαιρέτου ἐν τῇ ταινίᾳ Παθύρεως λεγομένης Πικρῶ Κόητιος τὸ ἐπιβάλλον αὐτῆ μέρος οὗ γείτονες νότου γῆ Σλήπτιος, βορρᾶ γῆ Πανᾶτος, ἀπηλιώ(του) ποταμὸς, λιβὸς γῆ Ἀραμήτιος ἢ οἱ ἄν ὡσι γείτονες πάντοθεν.

„Es hat verkauft Titos, die Tochter des Patus, Perserin, ungefähr 60 Jahre alt, mittelgroß, von gelbbrauner Hautfarbe, mit länglichem Gesicht und gerader Nase, mit ihrem älteren Sohn

Psenesis, Sohn des Harsiesis, Perserabkömmling, um 50 Jahre alt, mittelgroß, von gelbbrauner Hautfarbe, untersetzt, glatzköpfig, mit länglichem Gesicht und gerader Nase, am rechten Auge leidend, als Geschlechtsvormund von dem ihr und ihren Brüdern gehörenden, getreidebestandenen, ungeteilten Land, Pkro Koetios genannt, auf der Landzunge von Pathyris den auf sie entfallenden Anteil, dem benachbart sind im Süden das Land des Slepis, im Norden das Land des Panas, im Osten der Fluß, im Westen das Land des Harames oder wer die Nachbarn überall sein mögen.“³

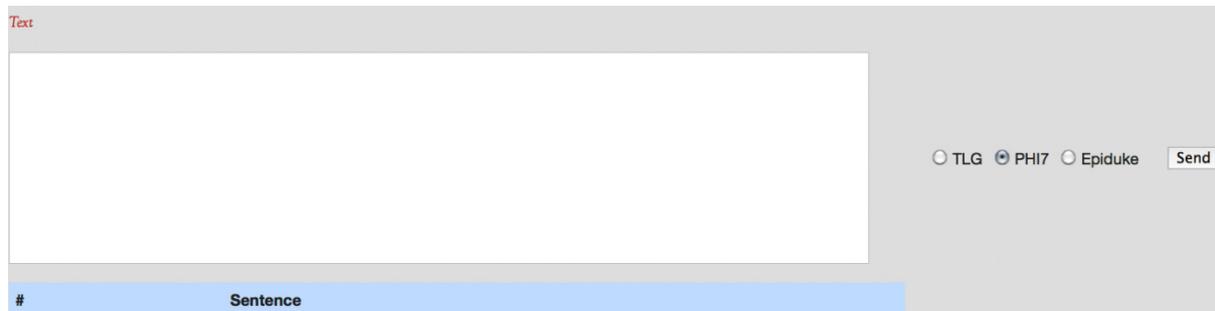


Abb. 3: Eingabemaske des Textergänzungstools

Das für den Test ausgewählte und fragmentierte Suchwort ist eines der Keywords des Kaufvertrages: ἀπέδοτο – es verkaufte. Das gesuchte Wort wird mit folgenden Ersatzzeichen in die Suchmaske eingegeben: Die fehlenden Buchstaben werden in eckige Klammern gesetzt und ihre Anzahl wird durch die entsprechende Setzung von Backslashes markiert (ἀ[\\]\\o).

Das Textergänzungstool bietet dem Nutzer die Auswahl aus unterschiedlichen Verfahren und ermöglicht so die Anwendung verschiedener Suchstrategien:

Kontext-Suche: Angezeigt werden Worte, die in der näheren Umgebung des gesuchten Wortes auftreten. Die Kontextsuche bietet sich als Verfahren an, wenn der Großteil des Papyrus und damit weitere Keywords lesbar sind.

Suche nach Worteigenschaften: Angezeigt werden Worte, die die gleiche Länge wie das gesuchte Wort aufweisen oder eine Buchstabenähnlichkeit.

Kombinationssuche: Der Nutzer kann aus verschiedenen Verfahren auswählen und damit die Suche stark präzisieren: Bigram, LetterNGrams, Named Entity, SentenceCooc, String Length, String Similarity.

³ Der vollständige Text des Papyrus wurde ediert und übersetzt von J. Hengstl (1978) Griechische Papyri aus Ägypten als Zeugnisse des öffentlichen und privaten Lebens. München: Heimeran Verlag unter der Nr. 142.

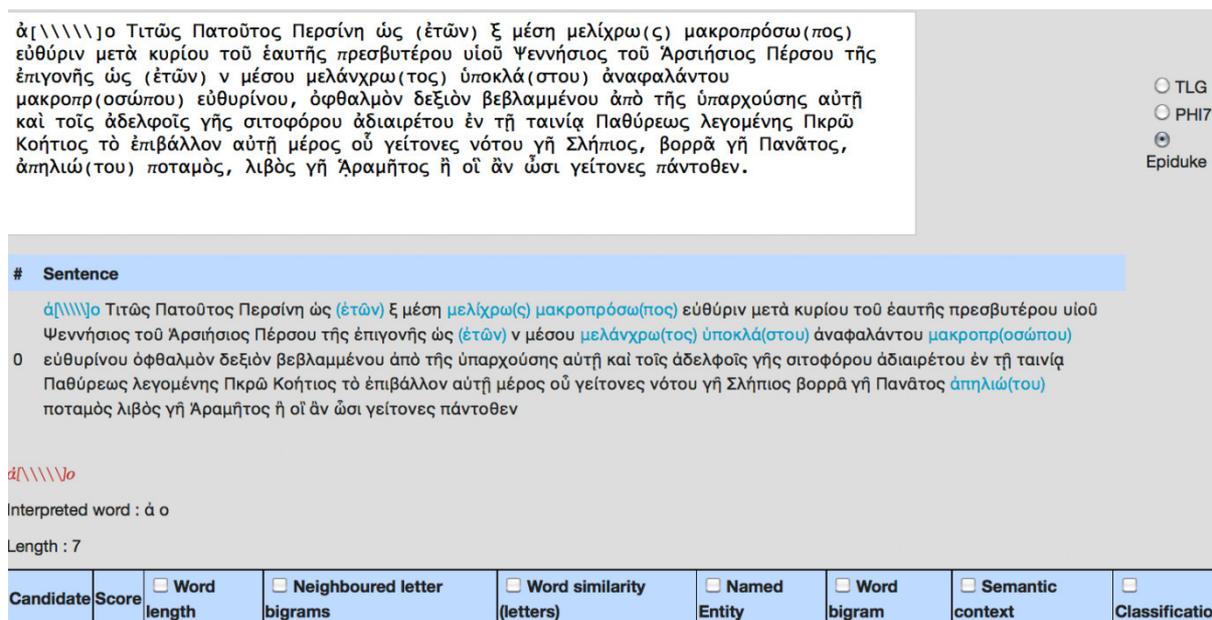


Abb. 4: Maske mit eingegebenem Text und Anzeige der einzelnen Verfahren

Die Vorstellung der einzelnen Verfahren erfolgt in der Reihenfolge der eingangs erwähnten Suchstrategien. Die beiden Verfahren, die zuerst vorgestellt werden gehören in den Bereich der Kontextsuche:

- a) *Bigram* (Syntaktisches Verfahren): Angezeigt werden Wörter, die den gleichen rechten oder linken Nachbarn wie das gesuchte Wort haben.
- b) *SentenceCooc* (Semantisches Verfahren): Es werden Wörter angezeigt, die normalerweise in Sätzen mit dem gesuchten Wort vorkommen.

Beide Verfahren nutzen die noch erhaltenen Umgebungswörter, um nach Wörtern mit ähnlichen syntaktischen oder semantischen Merkmalen zu suchen. Für den Beispieltext des Kaufvertrages sind folgende Vorüberlegungen in Bezug auf das zu erwartende Ergebnis relevant: Der Kauf selbst wird meist mit ὠνή bezeichnet. Das Formular des stilisierten Protokolls lautet in ptolemäischer Zeit vorwiegend: ἀπέδοτο (es verkaufte) + Name des Verkäufers – ἐπρίατο (es kaufte) + Name des Käufers + verschiedene Adjektive zur Personenbeschreibung, wie μέση (mittelgroß), μελίχρως (gelbbraune Hautfarbe), μακροπρόσωπος (längliches Gesicht) und εὐθύριν (gerade Nase).

Weitere Formulierungen, die man mittels einer Kontextsuche erwartet, sind: ὁμολογῶ πεπρακέναι καὶ ἀπέχειν τὴν τιμὴν („ich bestätige verkauft und Kaufpreis erhalten zu haben“).

á\|\|\|o
 Interpreted word : á o
 Length : 7

Candidate	Score	<input type="checkbox"/> Bigram	<input type="checkbox"/> LetterNGrams	<input type="checkbox"/> Named Entity	<input type="checkbox"/> SentenceCooc	<input type="checkbox"/> String Length	<input type="checkbox"/> String Similarity
Σοήρις	2	0.5			0.021739130434782608		
υἱός	2	0.5			0.043478260869565216		
ἀπέδοτο	2	0.5			0.5217391304347826		
ὀλοκλήρου	1				0.021739130434782608		
ἀργυροῦν	1				0.021739130434782608		
Ἄγαθος	1				0.021739130434782608		
ἐπρίατο	1				0.10869565217391304		
Νεχούτου	1				0.17391304347826086		
Ποπλίου	1				0.021739130434782608		
ἐκγόνους	1				0.021739130434782608		
πεντηκκοστής	1				0.021739130434782608		

Abb. 5: Ergebnisausschnitt der Kontext-Suche

Wie in Abb. 5 zu erkennen, findet sich in der Anzeige der Kontextsuche neben Personennamen auch eines der Keywords des Papyrus: ἐπρίατο (es kaufte). Der Nutzer hat nun die Möglichkeit das Ergebnis dieser Kontextsuche zu evaluieren. Dazu gibt man ἐπρίατο in die „einfache“ Suchmaske innerhalb des eAQUA-Portals ein, als zu durchsuchende Datenbanken sind EpiDuke⁵ oder PHI 7 DDP auszuwählen. Da die Textergänzung über die Daten von EpiDuke generiert wird, empfiehlt es sich auch für die Evaluierung diese Daten zu nehmen. Auch zeigt ein direkter Vergleich Unterschiede zwischen beiden Datenbanken auf: das Suchwort ἐπρίατο tritt bei EpiDuke 181-mal auf, bei PHI 7 DDP nur 150-mal. Für die Evaluierung ist vor allem die Anzeige der signifikanten Kookkurrenzen⁶ interessant. Ein Vergleich dieser mit den Keywords des Test-Papyrus zeigt folgendes Ergebnis:

Significant cooccurrences of ἐπρίατο

χαλκοῦ (88); ἐτῶν (73); μελίχρως (38); ὤς (75); εὐθύριν (28); μέσος (28); μακροπρόσωπος (26); . (44); ταλάντων (26); ταλάντου (18); Πετειαρσεμθεῦς (13); δραχμῶν (32); Πανοβχούνιος (12); τετανός (13); Νεχούτου (12); παρὰ (54); Πετειαρσεμθεῦς (10); τῆς (59); Πέρσης (14); νομίσματος (20); Πανεβχούνιος (8); Παθύρεως (10); τιμῆς (23); δηναρίων (15); στρογγυλοπρόσωπος (9); τῶν (52); οἰκετεῖαν (6); οὐλή (19); διαθήκης (9); Περσίνη (9); νόσου (8); Τοτοέους (7); ἀπέδοτο (9); Πακοίβιος (7); ὑποκλαστός (6); χρήματά (5); Ὄξυρύγχων (14); ζυγοστατοῦντος (5); δοῦλον (8); Φαγώνιος (6); γένει (8); καλεῖται (6); κτᾶσθαι (6); δραπετικόν (4); ῥέμβον (4); ἐνός (7); Ζμανρέους (5); ἐπιγονῆς (12); Ὠρος (14); ἐκ (26); ὑπάρχουσαν (8); αὐτῆς (23); ὠνήν (8); πίστει (7); δοῦλης (7); πόλεως (20); κατὰ (23); δοῦλην (7); κελεύοντος (5); τοῦ (60); τετρημένον (4); ἀργυρίου (19); ἑαυτῆς (8); Σπεμμίνιος (4); παρέδωκεν (5); τε (17); τέλος (9); τινη (6); μέσου (5); βεβαιοῦντος (4); Νεχούθου (4); δισχιλίας (7); ἱερᾶς (8); ἀριστερόν (5); ὀνόματι (9); μέση (5); τάλαντα (13); εὐθύριν (4); κοράσιον (4); Πετεύριος (5); Μαρκουπολίτης (3); ὑπόκλαστος (3); ἦ (15); μακροπρόσωπου (4); Κοβαετῆσις (3); ἦν (10); ἀπ' (9); Β (8); χρᾶσθαι (5); δραχμᾶς (24); δύο (16); ὑγιῆ (4); μετὰ (17); Ἡφαιστᾶ (3); διατάγματος (4); Ἀλεξανδρεῦς (4); ἀπέσχευ (4); ἀνέπαφον (4); ἀγορᾶ (4); Σούχου (5);

Abb. 6: Anzeige der signifikanten Kookkurrenzen zu ἐπρίατο

Das evaluierte Wort ἐπρίατο erscheint 9mal zusammen mit dem fragmentierten Wort ἀπέδοτο in EpiDuke, in der Datenbank PHI 7 DDP dagegen nur 6-mal. Neben ἀπέδοτο fallen auch andere signifikante Kookkurrenzen ins Auge, wie μέσος (28), μελίχρως (38), μακροπρόσωπος (26) und εὐθύριν (28) – allesamt Adjektive zur Personenbeschrei-

⁵ In dieser Datenbank befinden sich die griechischen Texte aus der Duke Data Bank of Documentary Papyri ergänzt durch die Metadaten (u.a. Titel, Stichwörter zum Inhalt) aus dem Heidelberger Gesamtverzeichnis.

⁶ Als Kookkurrenz bezeichnet man das gemeinsame Auftreten von zwei Wortformen in einem lokalen Kontext, vgl. dazu ausführlich G. Heyer / U. Quasthoff / Th. Wittig (2008) Text Mining: Wissensrohstoff Text. Konzepte, Algorithmen, Ergebnisse. Herdecke / Bochum: W3L-Verlag (korrigierter Nachdruck), bes. 23.

bung - oder ὠνήν (8), als Verweis auf den Kaufvertrag. Diese Wörter sind auch in dem Test-Papyrus belegt. Die Zahlen in Klammern markieren das gemeinsame Auftreten mit dem evaluierten Wort ἐπρίατο.

Durch die Anzeige der entsprechenden Belegstellen des gemeinsamen Auftretens, kann sich der Nutzer einen Überblick über vergleichbare Papyri verschaffen und diese auch sofort zur Gegenüberstellung mit dem fragmentierten Test-Papyrus heranziehen.

Einen sehr guten visuellen Eindruck über die benachbarten Wörter vermittelt der Kookkurrenzgraph.

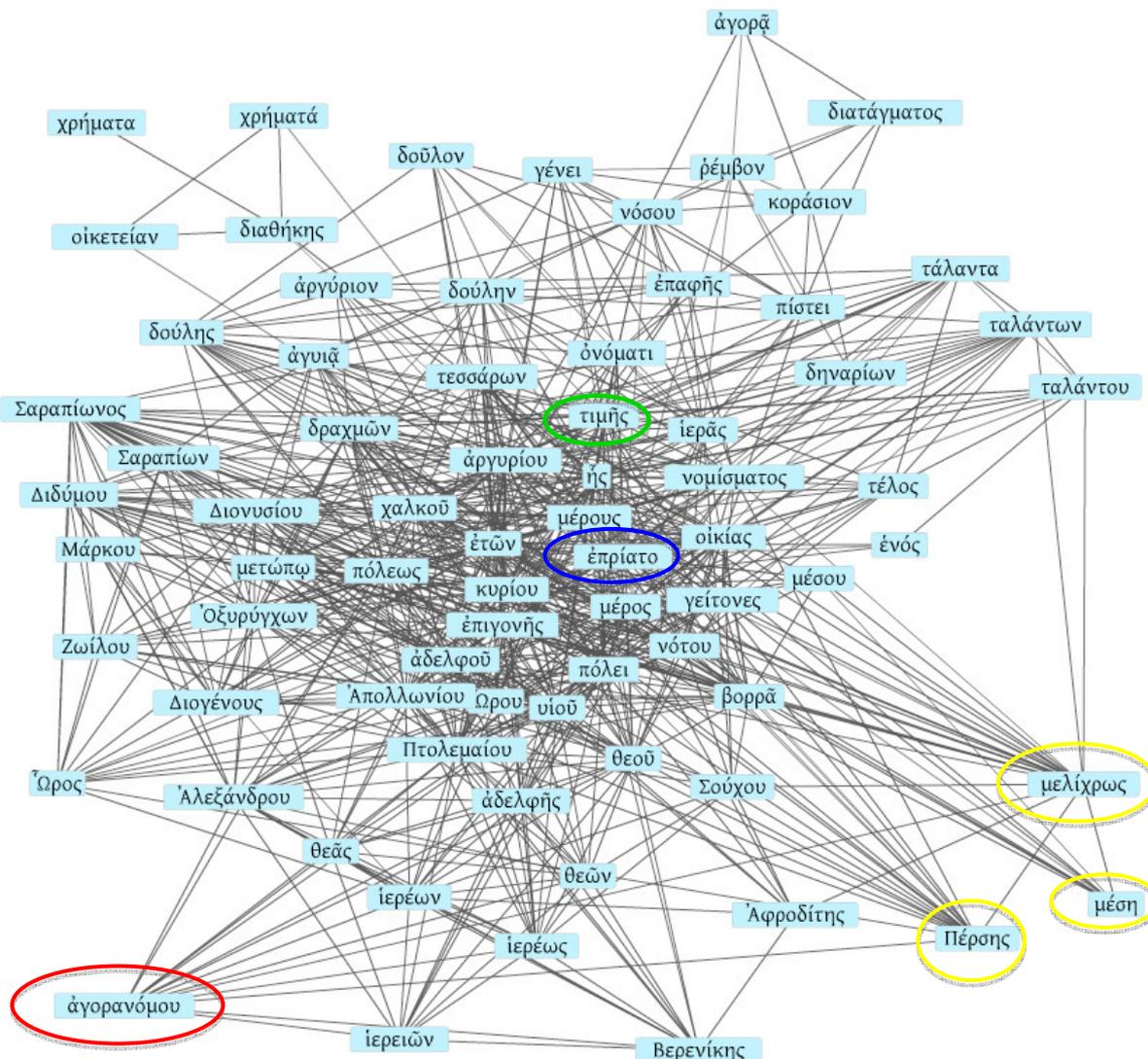


Abb. 7: Ausschnitt Kookkurrenzgraph für ἐπρίατο

Im Kookkurrenzgraphen sind ebenfalls neben verschiedenen Personennamen und den auch in der Liste der signifikanten Kookkurrenzen vertretenen Adjektiven zur Personenbeschreibung (gelb markiert) weitere relevante Keywords des Papyrus erkennbar, wie $\tau \iota \mu \eta \varsigma$ (grün markiert) oder der sehr häufig in Kaufverträgen in Erscheinung tretende Staatsnotar – ἀγορανόμος (rot markiert). Evaluiert man dieses Wort kommt man zu einem interessanten Ergebnis: Im Kontext der Kaufverträge erscheint das fragmentierte Wort ἀπέδοτο 65-mal zusammen mit ἀγορανόμος, davon 50-mal im Abstand von zwei Wörtern. In der Mehrzahl der Fälle steht ein Punkt zwischen beiden Wörtern.

Wie in Abb. 8 zu erkennen, schlägt das Tool Wörter vor, die – genau wie das gesuchte – mit α beginnen, mit o enden und eine Wortlänge von sieben Zeichen haben. Das fragmentierte Wort ἀπέδοτο steht mit einem Score von 2^9 an dritter Stelle.

Für die große Anzahl an fehlenden Buchstaben ist eine Textergänzung allein über die Wort-eigenschaften aber zu ungenau. Aus diesem Grund empfiehlt sich ein Verfahrencluster, eine Kombination der verschiedenen Suchstrategien: Bigram – LetterNGrams¹⁰ – SentenceCooc – String Length – String Similarity.

ἀ\\\\\\\\o
Interpreted word : ἀ ο
Length : 7

Candidate	Score	<input type="checkbox"/> Bigram	<input type="checkbox"/> LetterNGrams	<input type="checkbox"/> Named Entity	<input type="checkbox"/> SentenceCooc	<input type="checkbox"/> String Length	<input type="checkbox"/> String Similarity
ἀπέδοτο	5	0.5	3.0128546809192486E-5		0.5111111111111111	1.0	1.0
ἀποτίσο	3		1.8499049353136592E-5			1.0	1.0
ἀράσατο	3		4.309063665545054E-7			1.0	1.0
ἀδύνατο	3		1.5571988034933408E-7			1.0	1.0
ἀγομένο	3		7.751537752705404E-5			1.0	1.0
ἀμπαντο	3		5.550098309076602E-6			1.0	1.0
ἀνδραπο	3		2.8155842276429764E-4			1.0	1.0
ἀποδιδο	3		1.57984313704135E-5			1.0	1.0

Abb. 9: Ausschnitt der Ergebnisanzeige für die Kombinationssuche

Abb. 9 zeigt das fragmentierte Wort ἀπέδοτο als Ergänzungsvorschlag an erster Stelle mit einem Score von 5, d. h. alle Bedingungen der Suche sind erfüllt und die Ergänzung war erfolgreich.

Das Ergänzungstool bietet noch eine weitere Möglichkeit der Suche an, diejenige nach Named Entities. Darunter versteht man Wörter, die derselben Named Entity Kategorie – Personennamen, Ortsnamen, etc. – angehören. Als alleiniges Verfahren ist es jedoch nicht sinnvoll, hier empfiehlt sich ebenfalls ein Verfahrencluster aus Named Entities, String Length, String Similarity und den semantischen Kookurrenzen.

Πα\\\\\\\\ος
Interpreted word : Πα ος
Length : 8

Candidate	Score	<input type="checkbox"/> Bigram	<input type="checkbox"/> LetterNGrams	<input type="checkbox"/> Named Entity	<input type="checkbox"/> SentenceCooc	<input type="checkbox"/> String Length	<input type="checkbox"/> String Similarity
Πατούτος	5	1.0	1.3801961253595224E-4		0.6590909090909091	1.0	1.0
Παρίτιος	4		2.0120673128050492E-5	1.0		1.0	1.0
Πανιέτος	4		3.926265528948253E-6	1.0		1.0	1.0
Παμίνιος	4		8.84768352685093E-6	1.0		1.0	1.0
Παπίνιος	4		2.5230759648049618E-5	1.0		1.0	1.0

Abb. 10: Ergebnisausschnitt der Suche im Clusterverfahren mit Named Entities.

Das fragmentierte Suchwort für die Clustersuche in den Named Entities war Πατούτος (Πα\\\\\\\\ος). Die Ergebnisanzeige führt es an erster Stelle mit einem Score von 5, die Ergänzung war somit erfolgreich.

Textidentifizierung

Die Textidentifizierung bzw. Eingrenzung der Autorenschaft bei bisher nicht identifizierten literarischen Texten auf Papyrus, den so genannten "Wartetexten", soll in den nächsten Monaten mit Hilfe der Differenzanalyse und des Zitationsgraphen in Angriff genommen werden.

⁹ Die Bedingungen von String Length und String Similarity sind erfüllt.

¹⁰ Ein n-Gramm besteht aus n aufeinanderfolgenden Wortformen bzw. Buchstaben, vgl. G. Heyer (2008) 303.

Zusammenfassung

Die Arbeit des Teilprojektes will sich im Folgenden vor allem mit weiteren Tests des Textergänzungstools befassen. Diese sollen auf bisher unpublizierte Papyri ausgedehnt werden, so dass im Ergebnis auf der Basis von eAQUA Publikationen erfolgen können. Die Vorteile der automatischen Textergänzung lassen sich in zwei wichtigen Punkten zusammenfassen: Zum einen ermöglicht sie eine schnelle Bearbeitung großer Datenmengen – hier besteht auch die Möglichkeit der Auswahl des Referenzcorpus –, zum anderen kann der Nutzer aus verschiedenen Suchstrategien wählen oder diese für ein präziseres Ergebnis auch miteinander kombinieren. Nach der Einarbeitung der Klassifizierung in das Ergänzungstool erwarten wir eine Präzisierung der Suchanfrage, was sich auch auf die Vorschläge zur Ergänzung auswirken soll.

Schließlich kann für die Textergänzung auch das eAQUA-Suchtool herangezogen werden, wenn es darum geht, die Ergebnisse der Kontextsuche zu evaluieren. Die Anzeige der Belegstellen ermöglicht dem Nutzer eine sofortige Überprüfung der Ergebnisanzeige und damit auch ein nachhaltiges wissenschaftliches Arbeiten.

Michaela Rücker

Lehrstuhl für Alte Geschichte, Historisches Seminar, Universität Leipzig

mruecker1@me.com

Literatur

Heyer, G. / Quasthoff, U. / Wittig, Th. (2008) *Text Mining: Wissensrohstoff Text. Konzepte, Algorithmen, Ergebnisse*. Herdecke / Bochum: W3L-Verlag (korrigierter Nachdruck).

Pringsheim, F. (1950) *The Greek Law of Sale*. Weimar: Böhlau.

Rupprecht, H.-A. (1994) *Kleine Einführung in die Papyruskunde*. Wissenschaftliche Buchgesellschaft Darmstadt.

Wolff, H.J. (1978) *Das Recht der griechischen Papyri Ägyptens in der Zeit der Ptolemaeer und des Prinzipats*. 2. Band. München: Beck.