

DIGITAL RENAISSANCE

OLD CULTURES, NEW METHODS

Dr Thomas Köntges / @ThomasKoentges

Digital Humanities, University of Leipzig

STRUCTURE OF THE TALK

1. Digital Renaissance – What does this mean?
2. Prof Crane and his team
3. Perseus Digital Library and OPP
4. Showcasing some tools and methods
5. Questions

RENAISSANCE

- Carolingian Renaissance
- Italian Renaissance
- Digital Renaissance

Access to literature, education and reuse of knowledge

DIGITAL RENAISSANCE

- Increase proficiency in historical languages among as many people as possible
- Involving as many people as possible in the creation and restoration of knowledge
- Distribute restored/reconstructed and newly generated knowledge as widely as possible

DIGITAL RENAISSANCE

J. McGann (TLS, 20 November 2009):

“In the next fifty years the entirety of our inherited archive of cultural works will have to be re-edited within a network of digital storage, access, and dissemination. This system, which is already under development, is transnational and transcultural.”

PROF GREGORY CRANE

- PhD in Classics (Harvard)
- Editor-in-Chief Perseus Digital Library (Tufts)
- Alexander-von-Humboldt Professor (Leipzig)

HIS TEAM

- Leipzig Team
- Perseus Team

PERSEUS DIGITAL LIBRARY AND OPP

PERSEUS DIGITAL LIBRARY

OPEN PHILOLOGY PROJECT (OPP)

SHOWCASING SOME TOOLS AND METHODS

TEXTUAL DATA VS. ITS VISUALISATION

```
<tei.2>
<text>

<div1 type="Book" n="1" org="uniform" sample="complete">
<p>
<milestone n="1" unit="section">
Gallia est omnis divisa in partes tres, quarum unam incolunt Belgae, alia
</milestone></p>
</div1>

</text>
</tei.2>
```

ROBUST PERMANENT CITATIONS

```
urn:cts:greekLit:tlg0012.tlg001:1.10-1.20@οὐλομένην  
urn:cite:hmt:msB.1r  
urn:cite:hmt:chsmg.VA012RN-0013:0.0513,0.2216,0.12,0.0883
```

Using CTS/CITE URNs and a Catalogue

- [Perseus Catalogue](#)

MAKING TEXTS OPENLY ACCESSIBLE

- [Perseus Digital Library](#)
- [OPP Github Repositories](#)
- [Third Party Contributions](#)

OCR

booktif/0016/01000e.bin.png

stularemus misit in faciem Ascylti tunicā, & libe

stulatemus misit in faciemu Afelti tunicā, & libo

booktif/0016/01000f.bin.png

ratos querela iussit pallium deponere quod fo-

tatos querela iussit pallium deponete quod fo-

booktif/0016/010010.bin.png

lum hitem faciebat, & recuperato ut putabamus

lumu fitem ffaciebat, & recupetato lt putabamus

booktif/0016/010011.bin.png

thesauro in diuersorium præcipites abimus, præ-

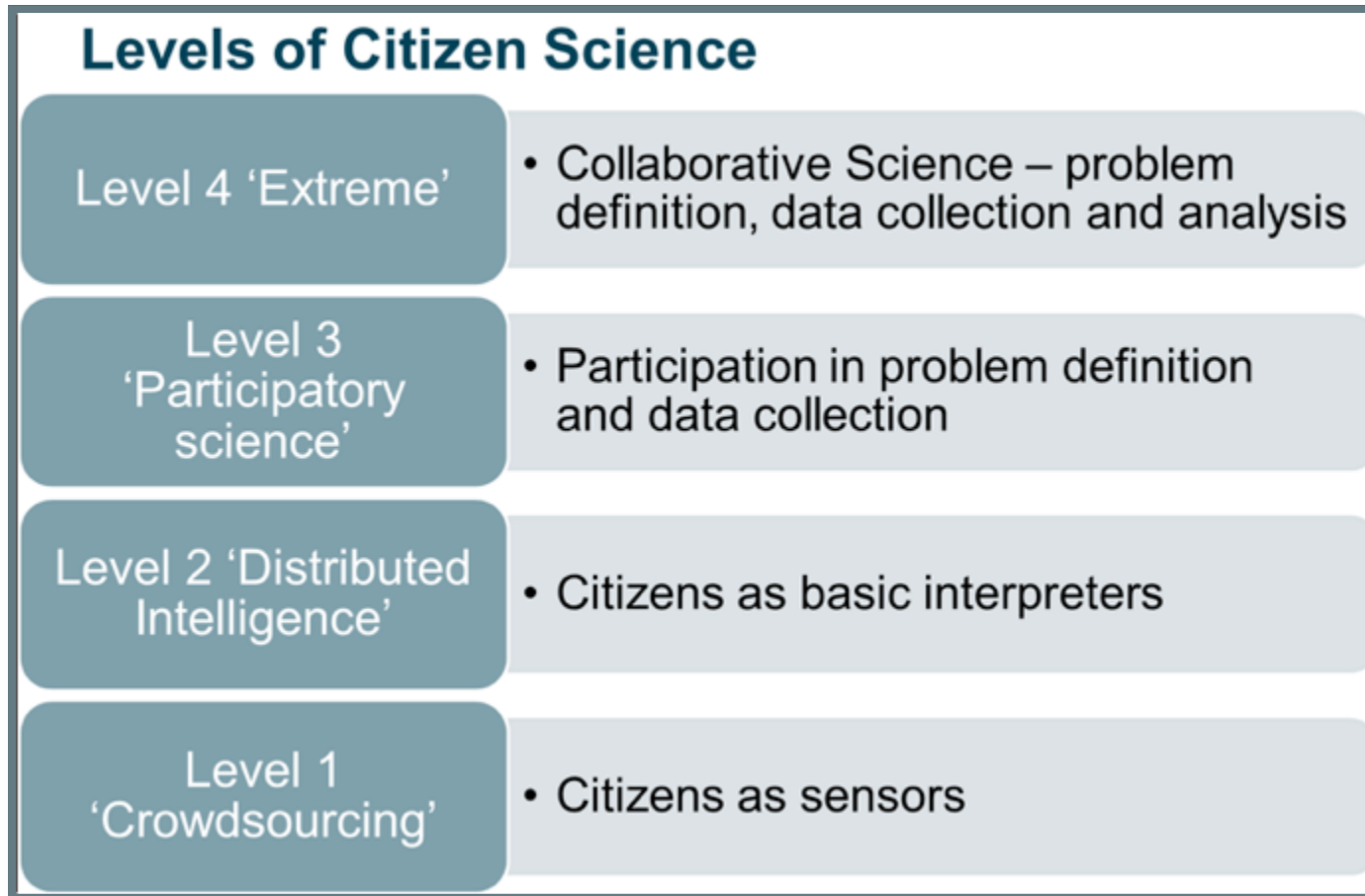
thesauro tn diuetforium ptrcipites abimus, præ-

booktif/0016/010012.bin.png

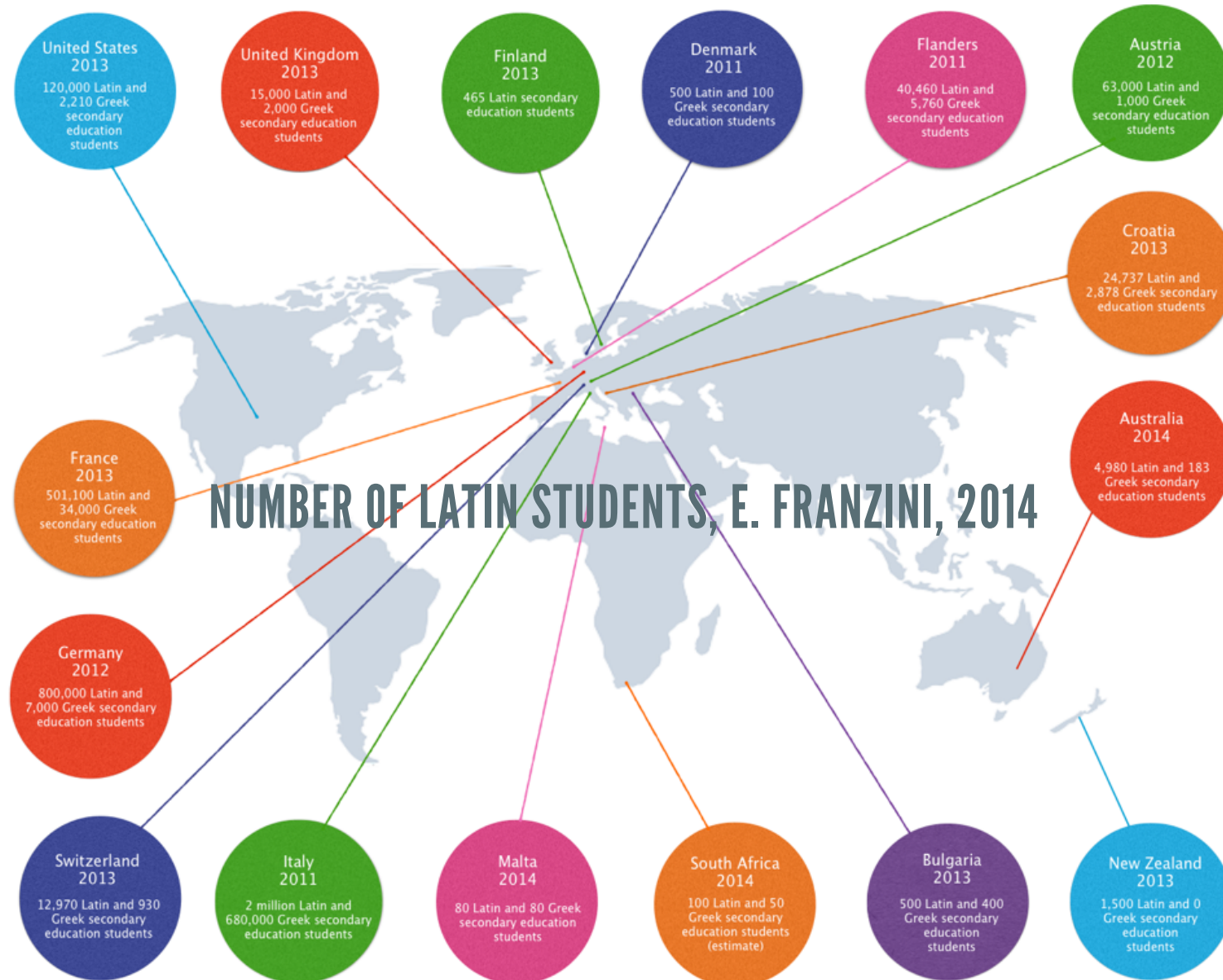
clusisq; foribus ridere acumē non minus concio

clufisqp foribus ridere acumē non minus concio

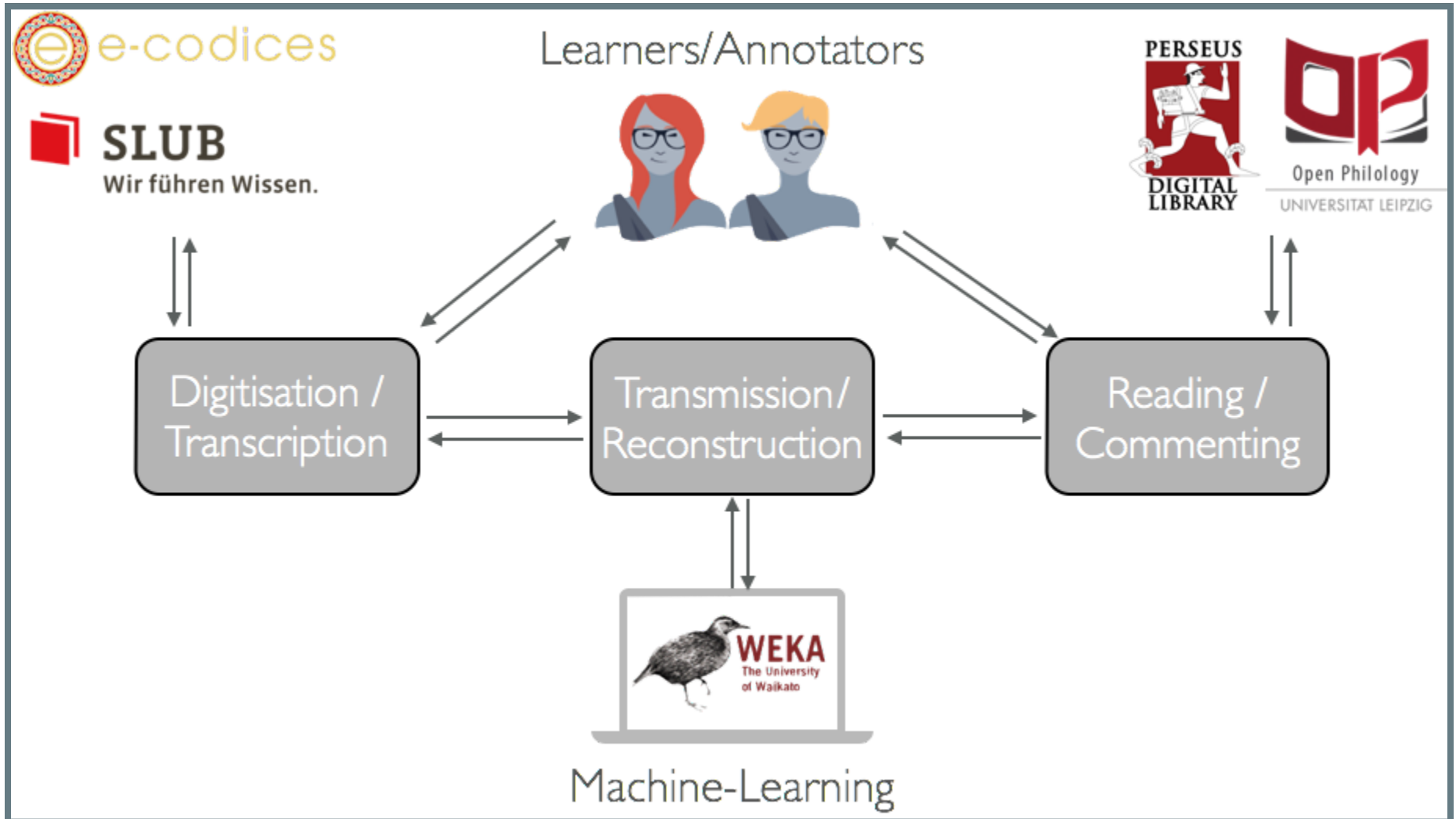
CITIZEN SCIENCE



Sui, D.Z., Elwood, S. and M.F. Goodchild (eds.), 2013.
Crowdsourcing Geographic Knowledge. Berlin: Springer



ARXE.TYPE WORKFLOW



INVOLVING EVERYBODY

SF_Handschrift2.txt

PETRONIUS ARBIER

CUM ALIO GENERE FURIA

rum declamatores inqui-
etantur; qui clamant haec
uulnera pro libertate pub-
lica excipi. hunc oculum
pro uobis impendi date
mihi ducem qui me du-
cat.ad liberos meos.na

succisi poplices. membra non sustinent. Hec ipsa
tollerabilia essent: siad eloquentia ituris uia
facerent. Nunc et rerv tumore et sententiary
uanissimo strepitu: hoc tantum proficiunt: ut
cum inforum ueniunt: putent se: in alium ter-
rarum orbem delatos. Etideo ego in adulescen-
tulos existimo in scolis stultissimos fieri.quia ni-
is que in usu habemus: aut audiunt
nt.Sed piratas cvm cathenis in litto-
et tyrannos edicta scribentes; q
ant filie ut capite n

SH_Handschrift2_verbessert.txt

petronius Arbiter

um.alio genere Furia
rum declamatores inqui-
etantur. qui clamant haec
uulnera pro libertate pub-
lica excipi. huncocolum
pro uobis impendi.date
mibi ducem quime du-
cat.ad liberos meos.na

succisi poplices. membra non sustinent.bec ipsa
tollerabilia essent: siad eloquentia ituris uia
facerent. Nunc et rerv tumore et senteniary
uanissimo strepitu: hoc tantum proficiunt: ut
cum inforum ueniunt: putent se.in alium ter-
rarum orbem delatos. Etideo ego in adulefcen-
tulos existimo in scolis stultissimos fieri.quia ni-
is que inufuhabemus: aut audiunt-
nt.Sed piratas cvm cathenis in litto-
et tyrannos edicta scribentes: q

	et
	m
	tur
	per

	nt
	quod
	esse
	pro
	est

ELEARNING BASED ON LINKED DATA (PROTOTYPE)


In Progress

Historical Languages eLearn

139.18.40.154:8000

Ancient Geek


Login to Ancient Geek!



Get your Greek on.


Learn Ancient Greek as you read Thucydides' harrowing tale of the war between
οἱ Λακεδαιμόνιοι (the Spartans) and οἱ Ἀθηναῖοι (the Athenians).

[Start Learning!](#)




Game your way through Greek.

Develop your mad Greek skillz through games and choose the path to
Greek glory that suits your personal learning style.



Keep it original.

Explore Classical texts in their original languages, the way they were
meant to be read, enhanced by data contributed by people like you.



Be a scholar!

Contribute micropublications like translations and annotations to help
create brand new, digital editions of Classical works!

ELEARNING BASED ON LINKED DATA

monica

فارسی ▼

Ancient Geek



زبان یونانی خود را ارتقا بده

همانطور که داستان‌های دلخراش توسیدید را درباره جنگ بین
οἱ Ἀθηναῖοι (آتنی‌ها) و οἱ Λακεδαιμόνιοι (اسپارته‌ها) شروع کن به یاد گرفتن!



پژوهشگر باش!

برای کمک به ایجاد نسخه‌های جدید و دیجیتال آثار کلاسیک، با کارهایی چون ترجمه و حاشیه‌نویسی به نشر میکرو یاری رسان!



فقط از اصل استفاده کن.

در متون کلاسیک به همان زبان اصلی کاوش کن، با همان زبانی این متن‌ها از همان اول نوشته شده بودند، متن‌هایی که به واسطه داده‌هایی که افرادی مثل تو در ایجاد آن همکاری کرده‌اند، بهبود یافته‌اند.



با بازی راه خود را در دنیای یونان پیدا کن.

مهارت‌های یونانی خاص خودت را از طریق بازی توسعه بده و راه خودت را به سمت شکوه یونان که متناسب با سبک یادگیری خودت است، انتخاب

ELEARNING BASED ON LINKED DATA

Lambda



Λ λ

Say it like the "l" in *l*ithography.


IPA [l] [Listen](#)

Lambda makes the same sound in **λίθος** (líthos), meaning "stone".

Think of *neolithic* (literally, new + stones), *lithology* (study of rock formation).

ELEARNING BASED ON LINKED DATA

[Home](#) [Learn](#) [Read](#) [Create](#)

χαίρε,  monica2 [Log out](#)

Πεντηκονταετία by Θουκυδίδης, 1.89.1

οι γὰρ Ἀθηναῖοι τρέψαν τοῦδε ἔλθον ἐπὶ τὰ πράγματα ἐν οἷς **ἡδύθησαν**.

Πεντηκονταετία by Θουκυδίδης, 1.89.2

ἐπειδὴ Μάδοι ἀναχώρησαν ἐκ τῆς Σαύρας
νυκθήμεναι καὶ ναυαὶ καὶ πεζοὶ ὑπὸ Ἑλλήνων καὶ
οἱ καταφυγόντες αὐτῶν ταῖς ναυσὶν εἰς Μυκάλην
διεφθάρησαν. Ἀνατυχίδης μὲν ὁ βασιλεὺς τῶν
Λακεδαιμονίων, ὅσπερ ἤγειτο τῶν ἐν Μυκάλῃ
Ἑλλήνων, ἀναχώρησεν ἐπ' οἶκον ἔχων τοὺς ἀπὸ
Πελοποννήσου συμμάχους, οἱ δὲ Ἀθηναῖοι καὶ οἱ

ἡδύθησαν from αὐτῶνσε [Parse](#)

Translations

English [Farsi](#)

grow in power

Other Resources


- [View Sentence as Tree](#)
- [Contribute Annotation](#)

Related Grammar Topics

- [None found](#)

119.18.40.154:8000/reader/#

ELEARNING BASED ON LINKED DATA

[Ancient Geek](#) [Learn](#) [Read](#) [Create](#) English ▾ 




ἡ χώρα,ας: The land, of the land

Case	Article, Noun
Singular	Plural
Nom.	ἡ χώρα
Gen.	τῆς χώρας
Dat.	τῇ χωρᾷ
Acc.	τὴν χώραν
Voc.	ὦ χώρα

Look out for Accents!

When a first declension noun has an acute accent on the last syllable in the nominative, the accent is changed to a circumflex in the genitive and dative, in both the singular and plural.

For example: ἡ στρατιά (meaning *the army*) will become τῆς στρατιᾶς in the genitive singular.



οἰκοδομία

Nominative, Singular, Feminine

Nominative, Singular, Feminine

οἶκο | I

δήλη ἡ ----- ἔτι καὶ νῦν ἐστὶν.

[More definitions for οἰκοδομία](#)

ELEARNING BASED ON LINKED DATA

14

Link

grc

οἱ γὰρ Ἀθηναῖοι τρόπῳ τοιῷδε ἦλθον ἐπὶ τὰ
πράγματα ἐν οἷς ηὐξήθησαν .

en

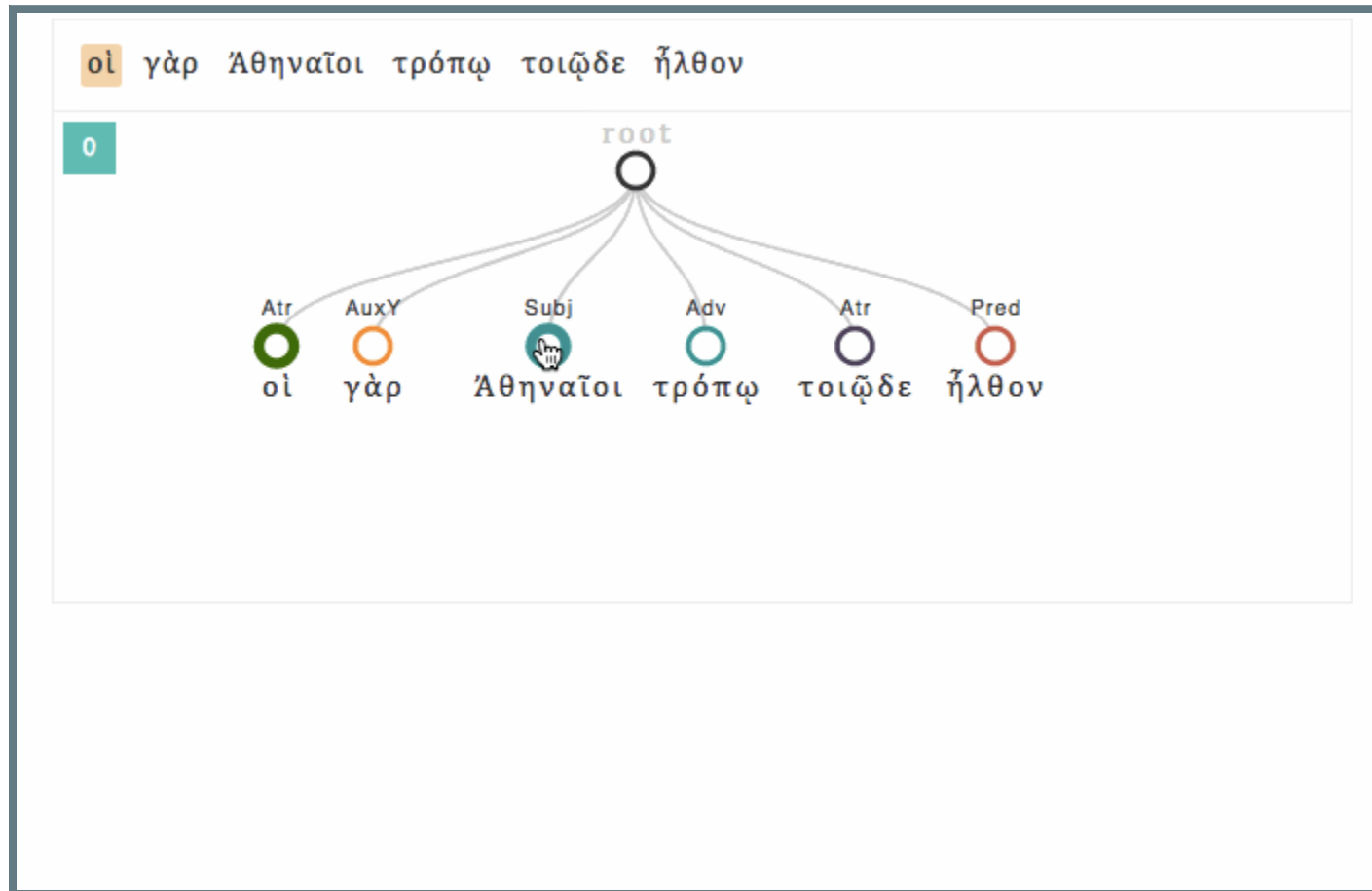
In this way the Athenians came to the circumstances

-

under

which they grew in power .

ELEARNING BASED ON LINKED DATA



COLLABORATIVE EDITING (WITH SCHOLARS)

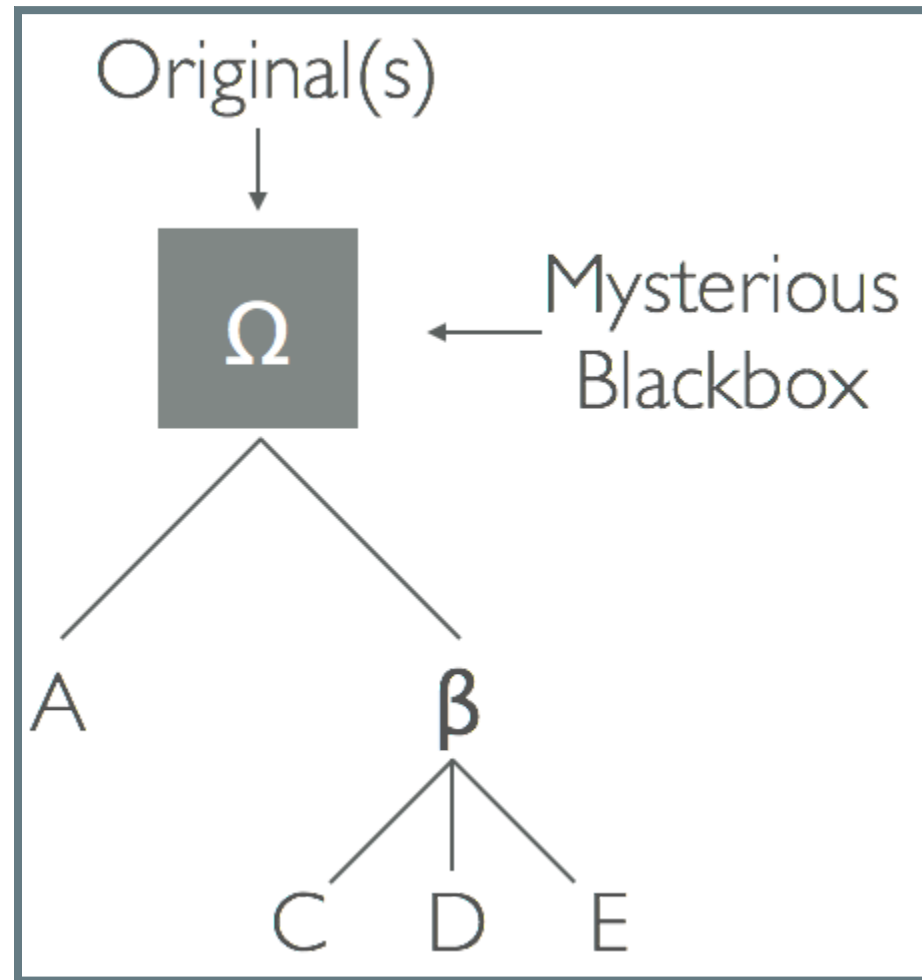
- Fragmentary Texts
- Thematic Annotations
- Treebanks

RESEARCHING TEXTUAL TRANSMISSION

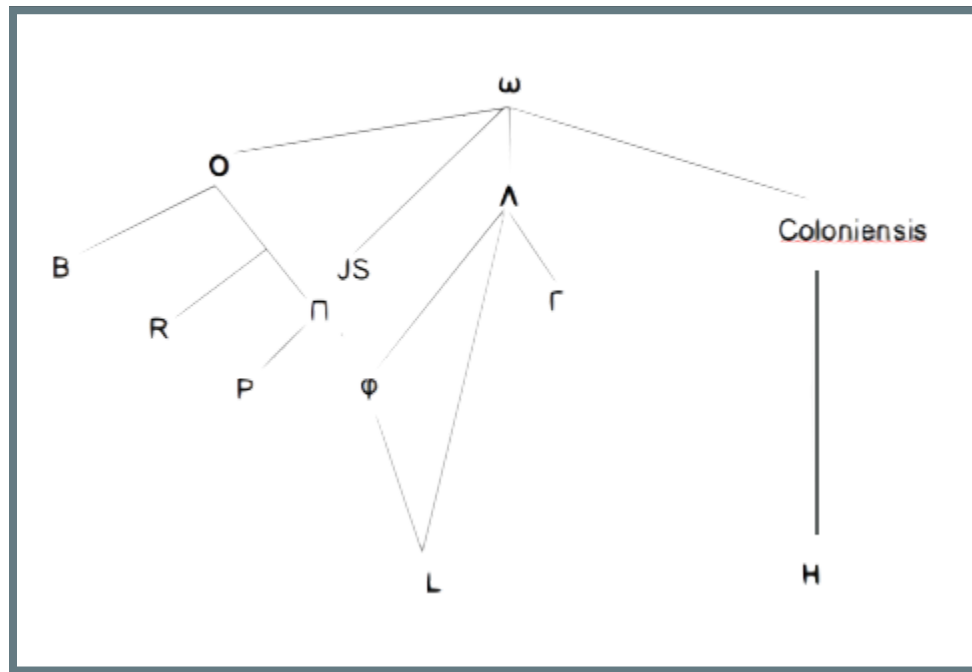
Textual transmission is a mono-directional communication process of text-data, graphics, and meta-data from the past to the present.



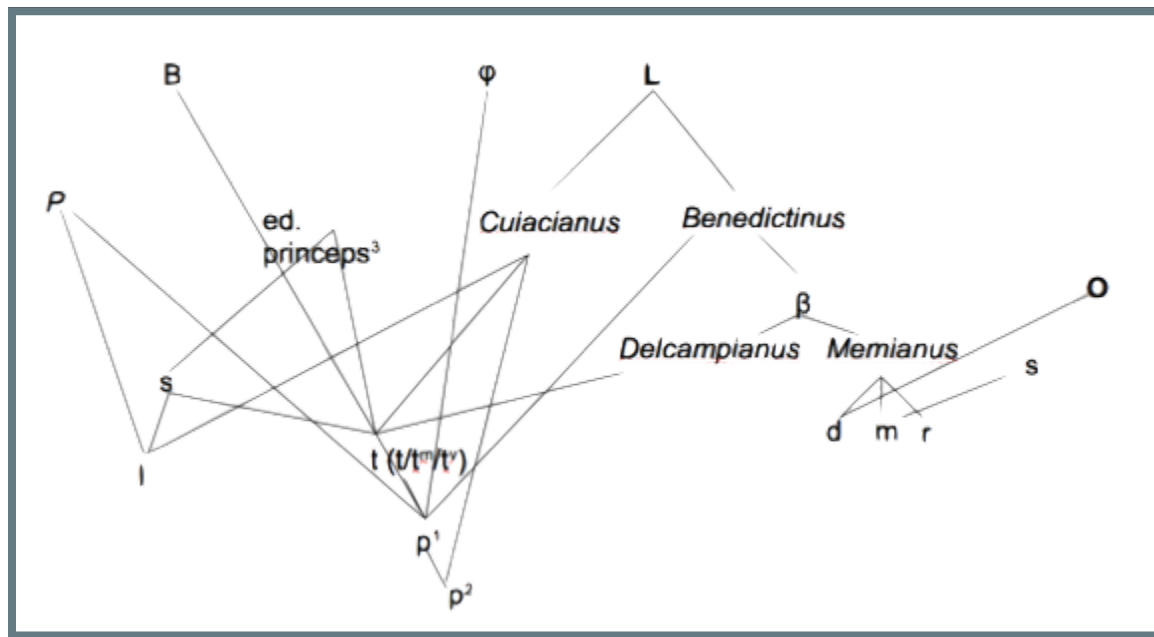
RESEARCHING TEXTUAL TRANSMISSION



RESEARCHING TEXTUAL TRANSMISSION



RESEARCHING TEXTUAL TRANSMISSION



- OPP Manuscript Catalogue

WEKA

- Weka 3.7.11: Open Source Data Mining Software in Java
- A Collection of machine-learning algorithms for data-mining tasks
- GUI, but also implemented in e.g R, Python
- Although Weka is a powerful tool, classifying the manuscripts will be a challenge

[Weka Homepage](#)

WEKA

Weka Explorer

Preprocess Classify Cluster Associate Select attributes Visualize

Classifier: AttributeSelectedClassifier -E "weka.attributeSelection.WrapperSubsetEval -B weka.classifiers.bayes.NaiveBayes -F 5 -T 0.01 -R 1 --" -S

Test options:

- ☐ Use training set
- ☐ Supplied test set Set...
- ☒ Cross-validation Folds 10
- ☐ Percentage split % 66

More options...

(Nom) Class

Start Stop

Result list (right-click for options):

- 17:45:58 - rules.ZeroR
- 17:46:08 - rules.OneR
- 17:46:21 - bayes.NaiveBayes
- 17:47:02 - lazy.IBk
- 17:47:23 - trees.J48
- 17:47:47 - trees.J48
- 17:49:36 - trees.DecisionStump
- 17:50:36 - trees.RandomForest
- 17:50:42 - trees.BFTree
- 17:50:48 - trees.FT
- 17:50:53 - trees.J48graft
- 17:51:01 - trees.LADTree
- 17:55:00 - trees.J48
- 17:55:27 - trees.RandomForest
- 17:55:40 - bayes.NaiveBayes
- 18:12:22 - trees.UserClassifier
- 18:21:47 - trees.J48
- 18:27:38 - bayes.NaiveBayes
- 18:41:15 - meta.AttributeSelectedClassifier

Classifier output:

Variant16

	7.0	10.0	1.0
quidem	7.0	10.0	1.0
om.	1.0	1.0	2.0
solum	3.0	2.0	1.0
[total]	11.0	13.0	4.0

Time taken to build model: 0 seconds

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	15	78.9474 %
Incorrectly Classified Instances	4	21.0526 %
Kappa statistic	0.6103	
Mean absolute error	0.2323	
Root mean squared error	0.3788	
Relative absolute error	61.4422 %	
Root relative squared error	87.8515 %	
Total Number of Instances	19	

=== Detailed Accuracy By Class ===

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
1	0.364	0.667	1	0.8	0.795	L	
0.7	0	1	0.7	0.824	0.833	O	
0	0	0	0	0	0.833	Phi	
Weighted Avg.	0.789	0.153	0.807	0.789	0.77	0.817	

=== Confusion Matrix ===

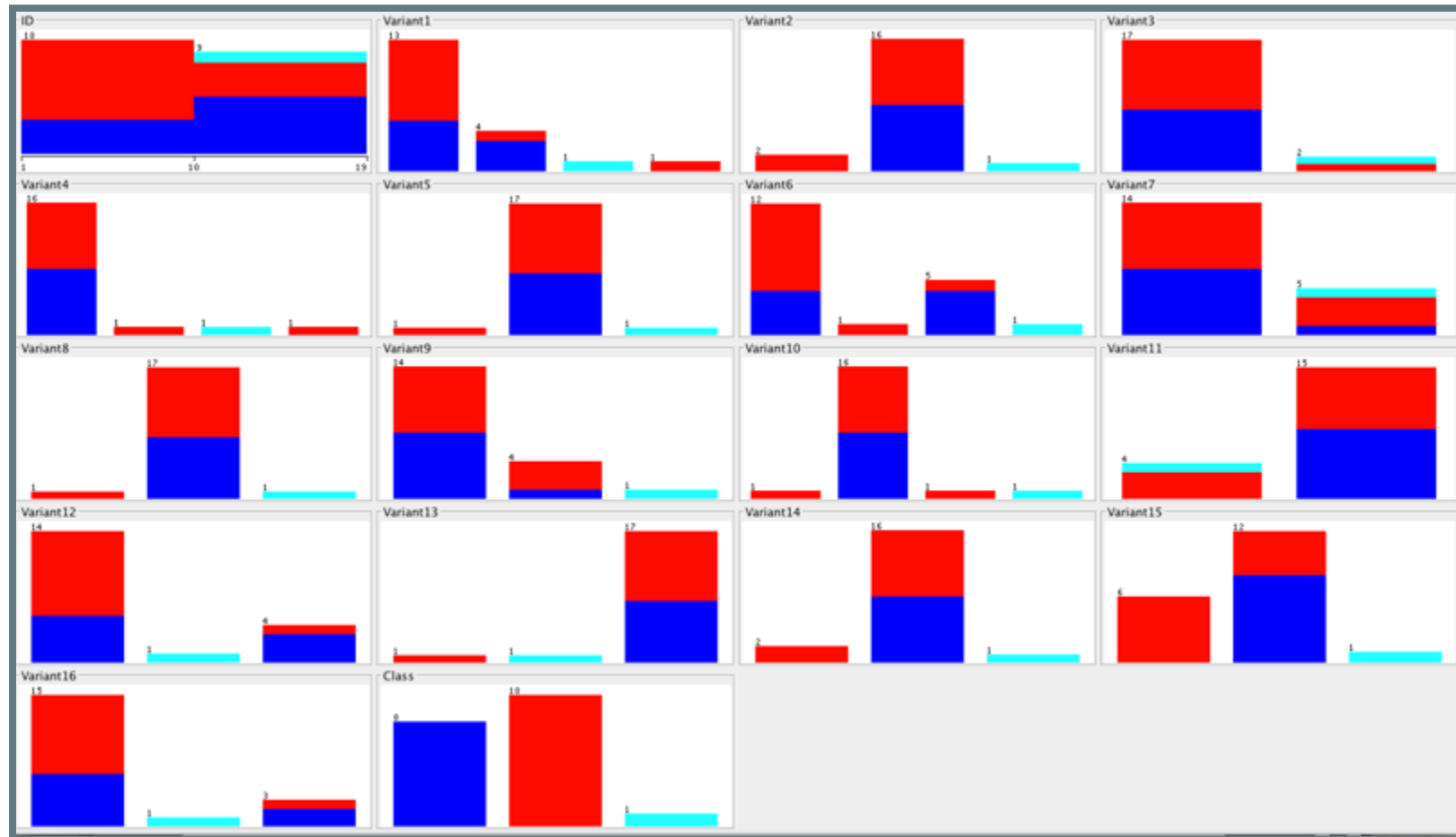
a b c <-- classified as

8	0	0	a = L
3	7	0	b = O
1	0	0	c = Phi

Status: OK

Log x 0

WEKA



WEKA ARFF FILE

```
@relation PetroniusManuscripts

@attribute ID numeric
@attribute Variant1 {Cum,Num,om.,Quum}
@attribute Variant2 {Nec,Non,om.}
@attribute Variant3 {Essent,om.}
@attribute Variant4 {ituris,nuris,om.,turis}
@attribute Variant5 {facent,facerent,om.}
@attribute Variant6 {rerum,runci,verborum,om.}
@attribute Variant7 {et,om.}
@attribute Variant8 {strepita,strepitum,om.}
@attribute Variant9 {terrarium_orbem,orbem_terrarum,om.}
@attribute Variant10 {qui,quia,quorum,om.}
@attribute Variant11 {his,iis}
@attribute Variant12 {et,om.,sed}
@attribute Variant13 {et,om.,sed}
@attribute Variant14 {pestilentia,pestilentiam,om.}
@attribute Variant15 {omnem,omnium,om.}
```

Structure of the ARFF Data file

DATA CLEANING, VISUALISATIONS, TOPIC MODELLING

- Open Refine (aka GoogleRefine)
- Google Fusion

BASIC D3 VISUALISATIONS

- Cartoonists' EQ Cartoons
- EQ Cartoons by Publisher
- How many Cartoons published on EQ
- Where were they published? (1)
- Where were they published? (2)
- Articles about EQ vs. Cartoons about EQ

TOPIC MODELLING

- Done by an algorithm
- Algorithm analyses chunks of texts
- Tries topics to find statistically significant word co-occurrences
- Evaluates the results and reruns
- Topics Index New Zealand (LDAvis)
- Topics Index New Zealand (FusionTable)

TOPIC MODELLING APP PROTOTYPE

- [Cartoon Topic Explorer](#)

QUESTIONS?

