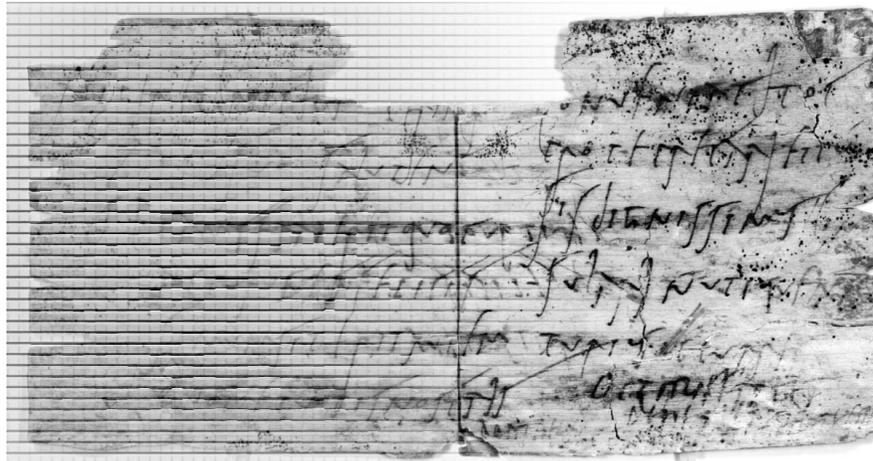


The Papyrologist's Assistant: Supporting the Reading of Ancient Texts



Dr Melissa Terras

Dept of Information Studies
University College London
m.terras@ucl.ac.uk

Henriette Roued-Cunliffe

Centre for the Study of Ancient Documents
University of Oxford
henriette.roued@classics.ox.ac.uk

Handwriting and Character Recognition

- Reading Handwriting is a primary aim of computing and engineering science
 - Vast research projects, various successes (OCR, etc)
 - Reading “difficult” texts beyond capacity of most computational approaches
 - Copperplate, dirty, noisy images, damaged, deteriorated
- What, if any, approaches can be used to assist papyrologists in reading damaged and abraded texts?
- Do we want them to computationally “read” them?
- What can we do to support – not replace – experts in reading texts?
- Case study regarding Vindolanda tablets

Vindolanda Texts

- Roman Fort on Hadrian's Wall, England
- Texts from AD 92 onwards
- Two types
 - ink texts
 - Carbon ink on wood. 300 texts survive
 - stylus tablets
 - recessed centre filled with wax. 100 texts
- Only contemporary and immediate written evidence of Roman Army in Britain



Vindolanda Ink Tablet



Vindolanda Tablet 291 Leaf No. 1 (front) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://vindolanda.csad.ox.ac.uk/4DLink2/4DACTION/WebRequestTablet?thisl

BBC NEWS BBC Undented Astro-C fd first direct BofS Barclays Popjustice Guardian eBay The Register

Vindolanda Tablets Online Tablets Exhibition Reference Help

Tablet 291 search results | new search

Image: Leaf No. 1 (front)	Latin text	English translation
<p>open image zooming viewer</p> <p>Vindolanda Inventory No. 85.057</p> <p>Introduction</p> <p>This diptych contains a letter to Sulpicia Lepidina from Claudia Severa, wife of Aelius Brocchus, sending Lepidina a warm invitation to visit her for her (Severa's) birthday (on the celebration of birthdays by private individuals see <i>RE</i> VII, 1142-4) and appending greetings to Cerialis from herself and greetings from her husband. The elegant script in which this letter is written is also probably to be recognised in 243, 244 and 248. The letters are slim, with marked ascenders and descenders, and very little use of ligature. There is occasional use of the apex mark for which see pp.57-61, above. In the present text the use is not always in long quantities. It is quite certain that the</p>	<p>i</p> <p>1 Cl(audia) · Seuerá Lepidinae [suae</p> <p>2 [sa]l[u]tem</p> <p>3 iii Idus Septembr[e]s soror ad diem</p> <p>4 sollemnem natalem meum rogó</p> <p>5 libenter faciás ut uenias</p> <p>6 ad nos iucundiozem mihi</p> <p>ii</p> <p>7 [diem] interuentú tuo facturá si</p> <p>8 [.].[c.3]s uacat</p> <p>9 Cerial[em t]uum salutá Aelius meus</p> <p>·[</p> <p>10 et filiulus salutant uacat</p> <p>11 m2uacat sperabo te soror</p> <p>12 uale soror anima</p> <p>13 mea ita ualeam</p> <p>14 karissima et haue</p> <p>Back</p> <p>15 m1 Sulpiciae Lepidinae</p> <p>16 Cerialis</p> <p>17 a S[e]uera</p>	<p>"Claudia Severa to her Lepidina greetings. On 11 September, sister, for the day of the celebration of my birthday, I give you a warm invitation to make sure that you come to us, to make the day more enjoyable for me by your arrival, if you are present (?). Give my greetings to your Cerialis. My Aelius and my little son send him (?) their greetings. (2nd hand) I shall expect you, sister. Farewell, sister, my dearest soul, as I hope to prosper, and hail. (Back, 1st hand) To Sulpicia Lepidina, wife of Cerialis, from Severa."</p>

Notes

[open notes viewer](#)

[TVII Abbreviations and Bibliography](#)

[addenda](#)

[print-friendly tablet view](#)

Done

Search for tablets - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://cairo.csad.ox.ac.uk/vto/index.php/tablets/search-for-tablets?tablet=128

Most Visited ... View all tablets ... Find Images ... Search fo... x ... Google Calen... AF The Binary He... uninvited gue... G2 | From the ... vindolanda ta...

Vindolanda Tablets Online II

Tablets Indices References About

127

TABLET 128

DESCRIPTION

Inv.no.85.242.a. 75 x 36 mm.

A fragment of a leaf which is complete at the right-hand edge and perhaps at the foot. It contains a report made by an optio named Arquittius in a form which is paralleled in [127](#) (see also above, p.76).

EDITION

.....
renuntia[ue]r[unt optio-
* nes et curatores
* detulit Arquittius optio
(centuriae) Crescentis

COMMENTARY

2. *curatores*: the term occurs as a rank or regular position in units with cavalry, see [Breeze \(1974\)](#), 282-3, [RMR](#) 47.i.7, [CEL](#) 82 (= [ChLA](#) X 431). Elsewhere, it probably denotes a specific function or task rather than a regular rank ([Gilliam \(1986\)](#), 109-13, [O.Flor.](#), p.24) and this may well be the case here. If we had the local equivalent of *curatores praesidii* (see [O.Flor.](#), loc. cit.), they might be in charge of small detachments outposted to local fortlets (cf. [154.16](#) note and introduction).

3. *Arquittius*: for this name as a gentilicium see [LE](#) 126, 403; it does not seem to occur as a cognomen.

4. For a century of *Crescens* see [148.2](#) and note. The name is very common.

129

View images

Done

Stylus Tablet - Wax Intact



Stylus Tablet - Wax Removed



Close up - Tablet 1563



- Complex incisions
- Woodgrain
- Surface discoloration
- Warping
- Cracking
- Noisy image
- Palimpsest
- Long process

What Is The Problem?

Need to build a system which **aids** in the transcription of the stylus texts

- Need to understand the process of reading an ancient text
- Information from the Vindolanda ink texts
 - Palaeographical
 - Linguistic
- Mobilise knowledge of these to implement a system
- Dovetail with Image Processing System
 - Cognitive Image Understanding System

Tackling the Problem

- Need to model process experts use as a basis for a computer model
- Need to build up a dataset of palaeographic and linguistic information to train a computer system, based on expert information
- Need to combine the model and the information in a system that will output *possible* and *plausible* interpretations

Modelling Expert Behaviour

- Modelling expert behaviour is a common approach used in Artificial Intelligence and Cognitive Psychology
- Two benefits
 - Modelling a process shows that you understand the process
 - Making an explicit model of the process provides the basis for the design of a computational system

The Papyrologist at Work

- Little research done into how papyrologists read and make sense of ancient texts
- Little research done on the process of reading damaged or ambiguous texts
- Little research done on the role of knowledge and reasoning in the analysis and understanding of complex images

Knowledge Elicitation

- Experts are notoriously bad at talking about their expertise
- Structured process for making explicit often unconsciously-mobilised knowledge used by an expert
- Developed protocols
 - Knowledge Library
 - Structured Interviews
 - Walk throughs
 - Transcripts
 - Analysis of discussions

Understanding the Papyrologists

- For Vindolanda
 - Two volumes of published ink texts
 - Possible to do computational analysis of published commentaries
 - Access to experts
 - Willing to be studied
 - Think Aloud Protocols
 - Knowledge Library

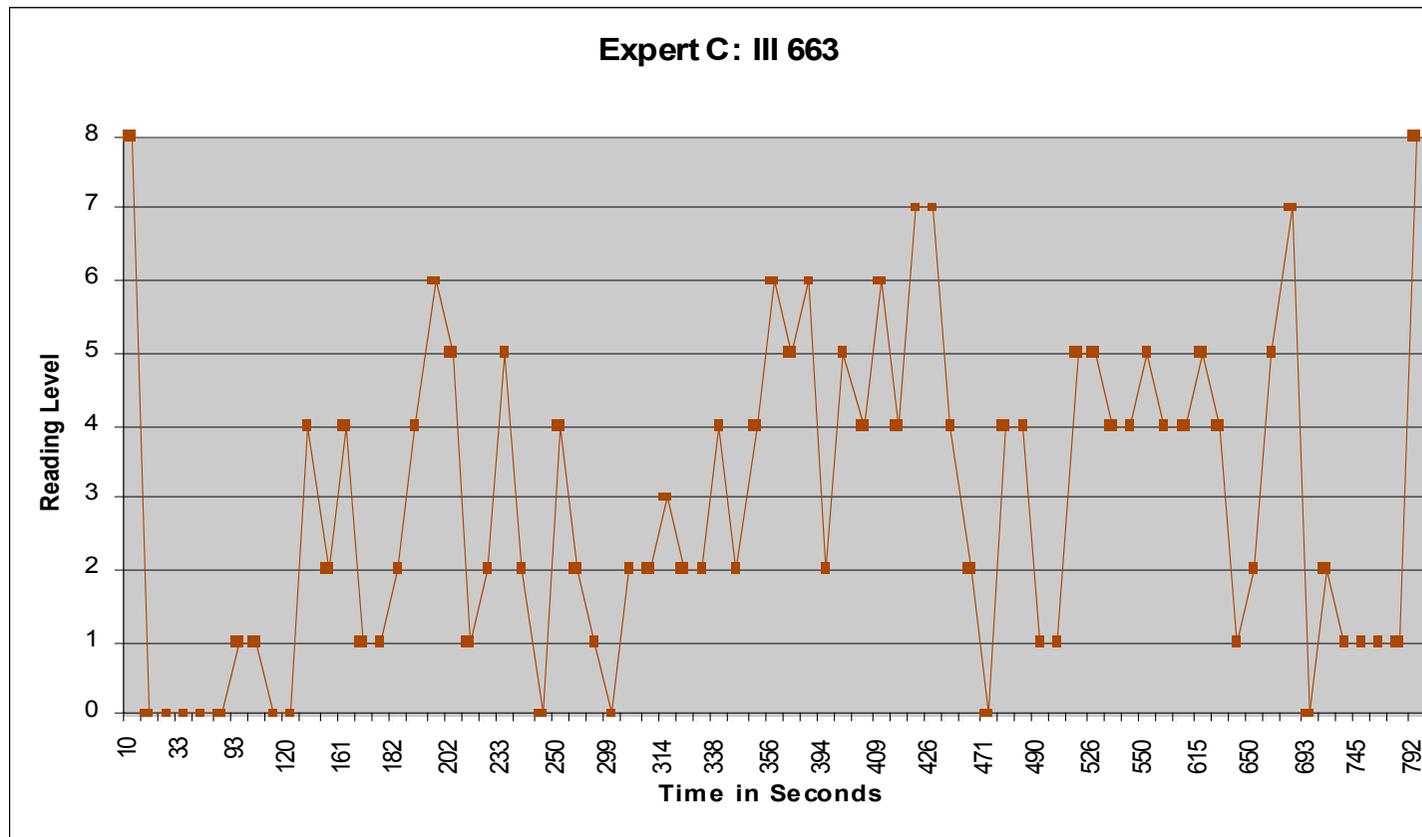
Basic Textual Analysis

- Using TACT and Wordsmith
- Allows analysis of the types of words used when discussing ancient texts
- Collocates
- Frequency
 - Ink Texts:
 - HORIZONTAL, BOLD, FORMAT, and DISCOLORATION, HYPOTHESIS, REASON
 - Stylus texts:
 - AFRAID, ASSUME, CONFUSING, CONVINCED, DECIDING, SURPRISED, and TRIED
- Analysis of the Latin itself
 - 10% of the characters in the published commentaries are marked as being uncertain

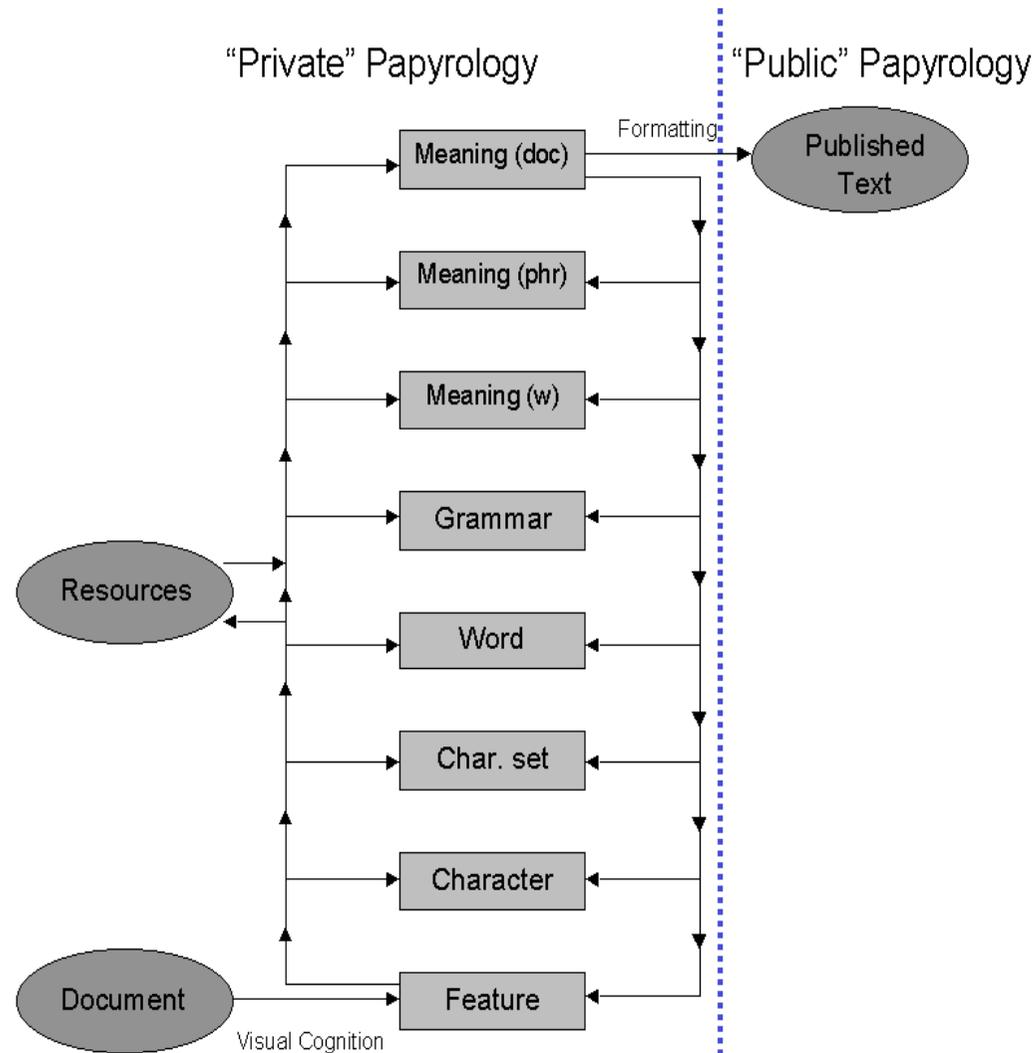
Encoding Scheme

Reading Level	Thematic Subject
8	Meaning or sense of document as a whole
7	Meaning or sense of a group or phrase or words
6	Meaning or sense of a word
5	Discussion of grammar
4	Identification of possible word or morphemic unit
3	Identification of sequence of characters
2	Identification of possible character
1	Discussion of features of character
0	Discussion of physical attributes of the document
-1	Archaeological or historical context

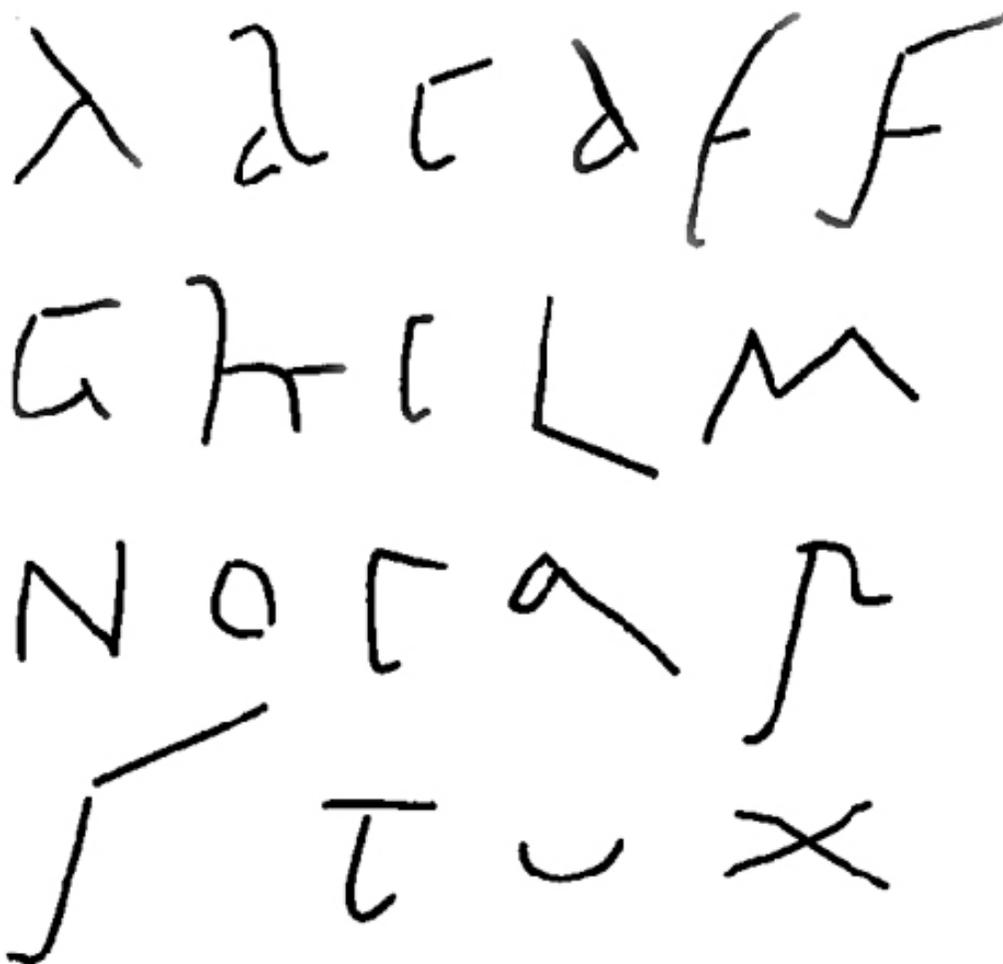
Content Analysis (2)



Model of the Papyrology Process



Palaeographical Information



- Old Roman Cursive (ORC)
- Every day Roman Script
- Same used on ink and stylus?
- Forensic evidence
- => ink info can be used for stylus texts

Collecting Linguistic Information

- Corpus of Vindolanda ink tablets
 - only contemporary linguistic resource
 - 300 texts plus fragments
- 6532 words, 2433 unique tokens, and 27364 characters
- Word list
- Letter frequency
- Information that can be incorporated into the system at the character and word level

Can computers ever read ancient texts?

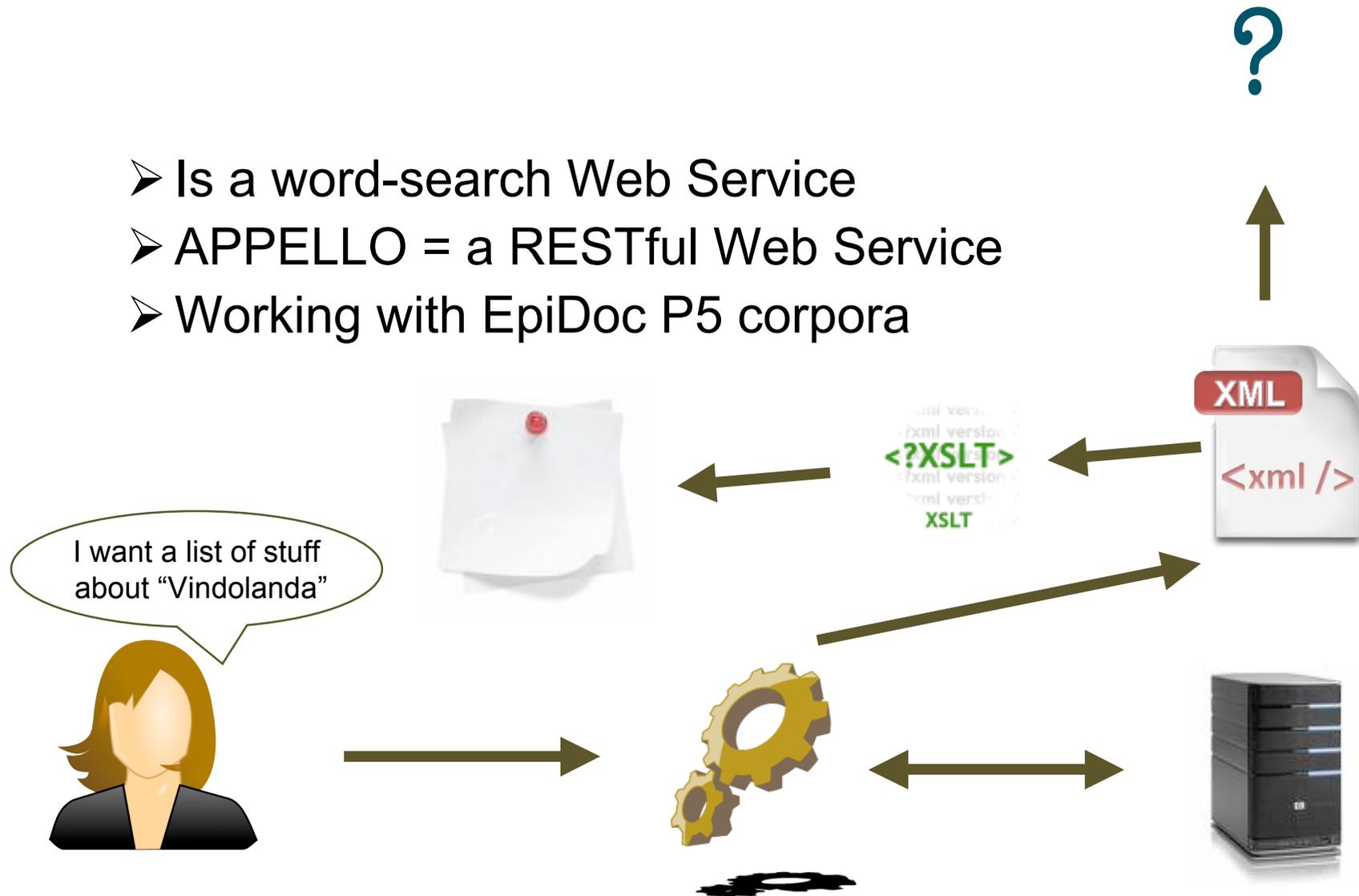
- Well, they can provide suggestions, based on known evidence
- They can keep a record of hypotheses encountered, discounted, and followed

Decision Support System for Reading Ancient Documents - doctoral research

- IT tools should do the jobs that humans find difficult:
 - Remembering complicated reasoning
 - Searching huge datasets
 - Accessing other experts knowledge
 - Enable cooperation between experts on a single document

APPELLO - searches huge datasets

- Is a word-search Web Service
- APPELLO = a RESTful Web Service
- Working with EpiDoc P5 corpora



➤ Methods:

- `get_tablets`
 - A list of all the tablets in the dataset
- `get_tablet`
 - Param = tabletID
 - Gets the tablet you've asked for
- `get_word`
 - Param = pattern
 - Get's all the words with a specific pattern

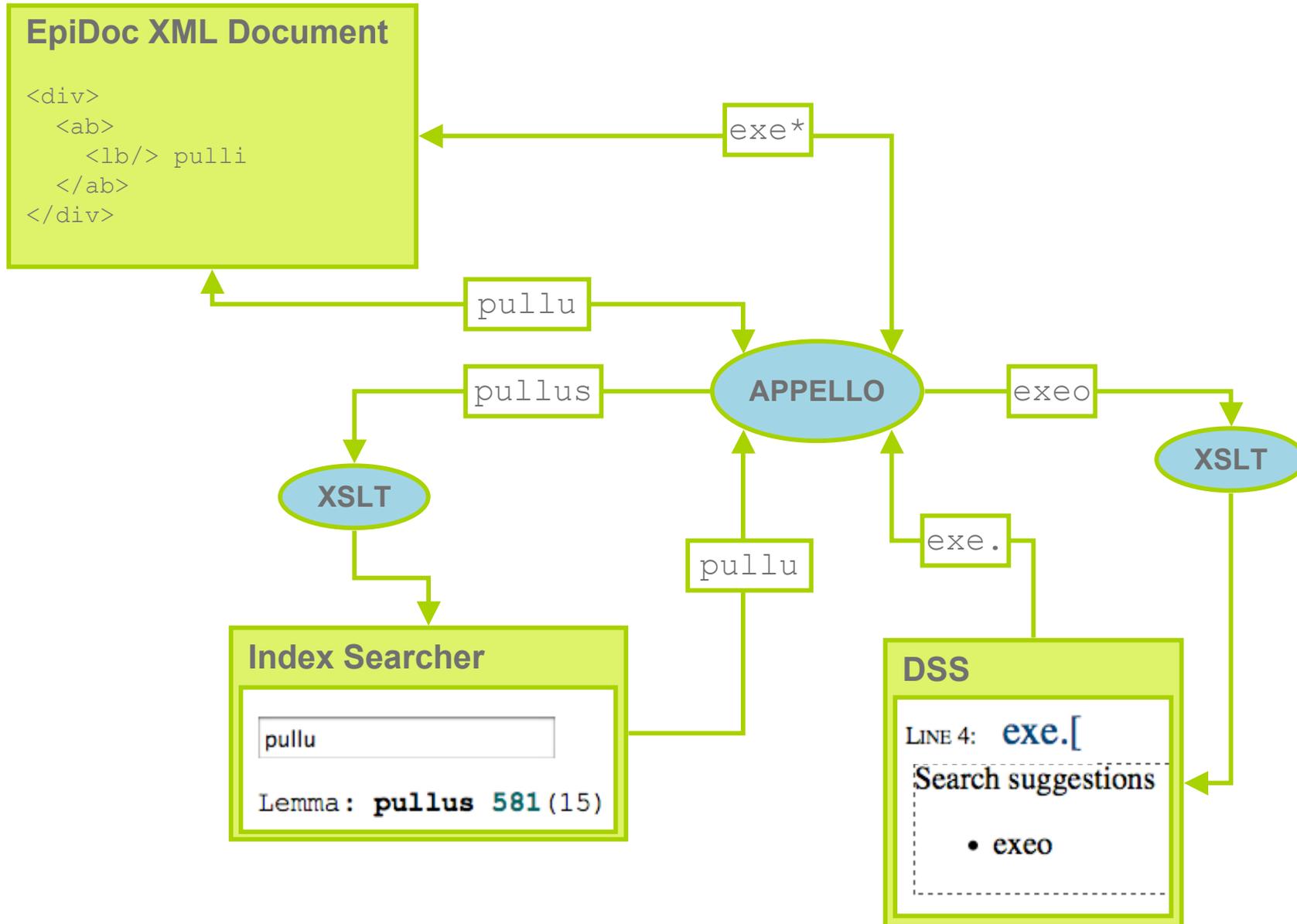
➤ Pattern:

Word: **ceruesa** 581 (1) 628 (1)

Word: **ceruesarius** 581 (2)

Word: **materies** 600 (1)

Person: **Nepos, Haterius** 611 (1)



- So we can build a more interactive website for the Vindolanda tablets
 - And add a word search facility to a Decision Support System

Word: **pullus** 581 (15) 582 (3) 616 (1)
 Word: **apud** 581 (2) 622 (1) 656 (1)
 Word: **ampulla** 589 (3)
 Word: **purpureus** 596 (1)
 Word: **caput** 613 (1)
 Word: **puto** 615 (1) 618 (1)
 putauī 629 (1)
 Word: **opus** 642 (1) 667 (1)



Word: **pullus** 581 (15) 582 (3) 616 (1)
 Word: **ampulla** 589 (3)

LINE 4:



CURRENT INTERPRETATIONS

Word: exe*

Search suggestions

- exeo

Next steps

- Use APPELLO more!
 - Via Google Gadgets
 - Via HTML add-on (to use from any blog/website)
 - From Decision Support Systems

- APPELLO is now working for Vindolanda tablets - try it out with:
 - Monumenta Asiae Minoris Antiqua XI
 - Inscriptions of Aphrodisias

- Develop a DSS prototype to show how it would work

To conclude

- Can Computers ever read ancient texts?
 - Maybe, but not in the near future
- Wrong question to ask:
- Can Computers ever be used to *aid* in reading ancient texts
 - Yes
 - Developing an understanding of how we can use technology to aid papyrologists brings an understanding of papyrology itself.

vto2.csad.ox.ac.uk

