



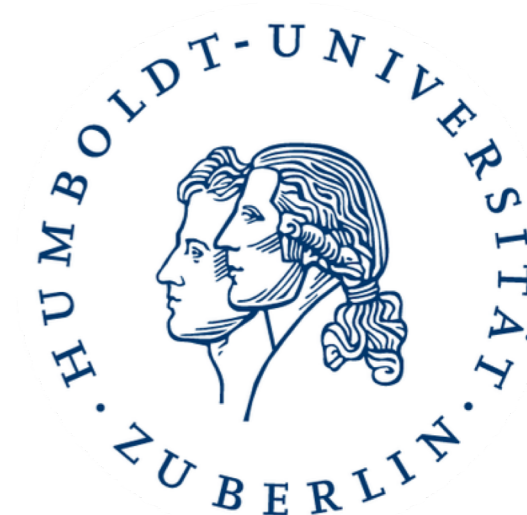
CALLIDUS

Computer-Aided Language Learning:
Lexikonerwerb im Lateinunterricht durch korpusgestützte Methoden

Humboldt-Universität zu Berlin

Computer and Media Service, Corpus Linguistics, Pedagogy of Classics

Andrea Beyer, Malte Dreyer, Stefan Kipf, Milena Kühnast, Anke Lüdeling, Konstantin Schulz, Andreas Vollmer



Research Questions

- Question 1: How can we apply available **corpora** of classical texts for the **acquisition of Latin** at school?
- Question 2: How can we systematically obtain **lexical information** from problematic texts suitable to enhance the acquisition of Latin **vocabulary**?
- Question 3: What are the criteria of a user-friendly **software** in a **teaching/learning** environment?
- Question 4: How can we model digital language **exercises** according to the requirements of a **cognitive progression** design?

Introduction

- **Current state** of vocabulary learning in high-school Latin classes:
 - **Infrequent** vocabulary-centered input
 - Methodologically **not adapted** to recent technical developments (digitization)
 - Usually based on core vocabularies that are **not scientifically recognized**

Aim

- Technical framework for **vocabulary learning** in Latin:
 - Flexible, individualized exercises **on a regular basis**
 - Designed for **mobile devices** and learning environments, but also printable
 - Dynamic **corpus-based** vocabulary filter

Method

Resource Workflow

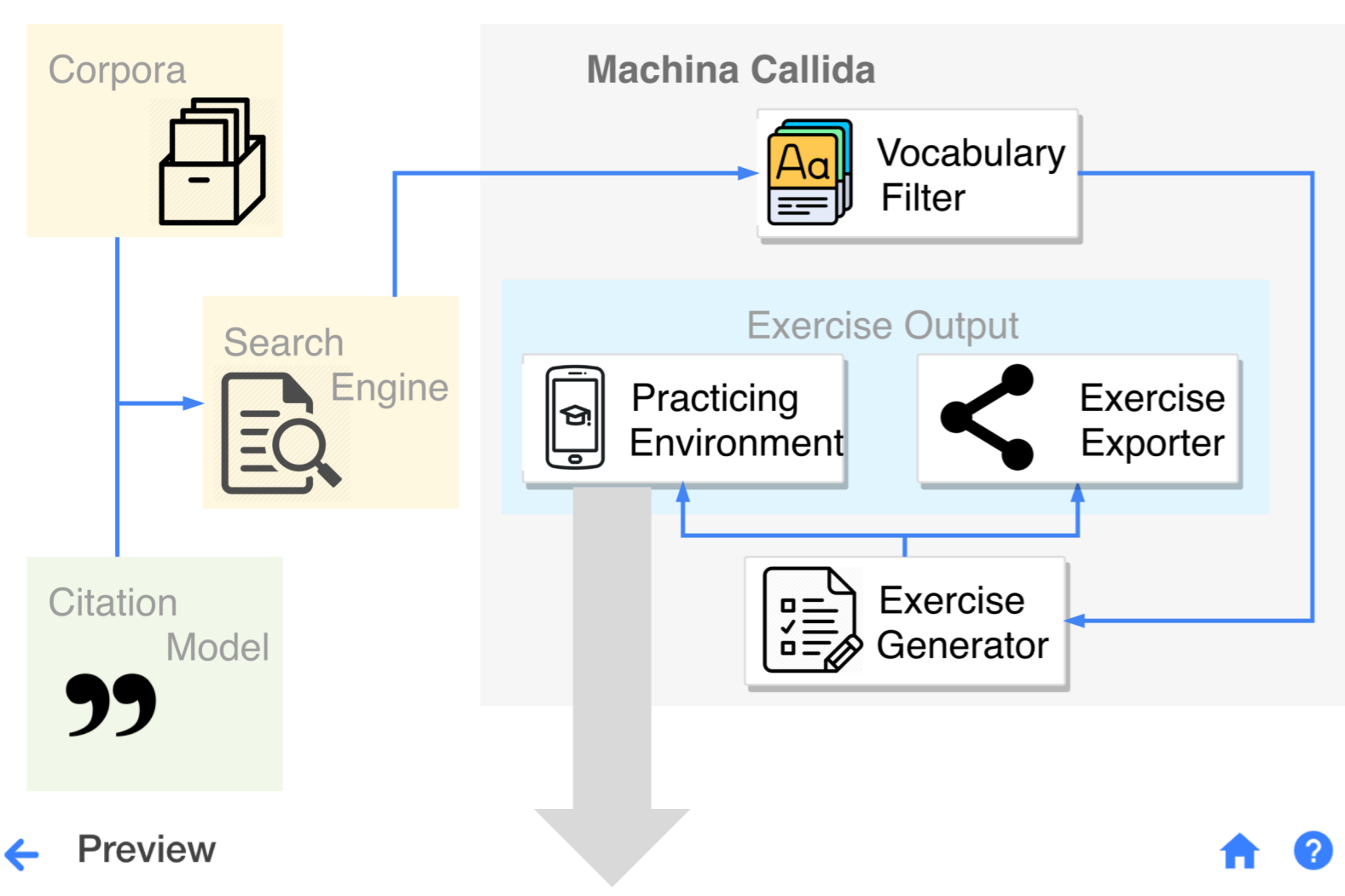
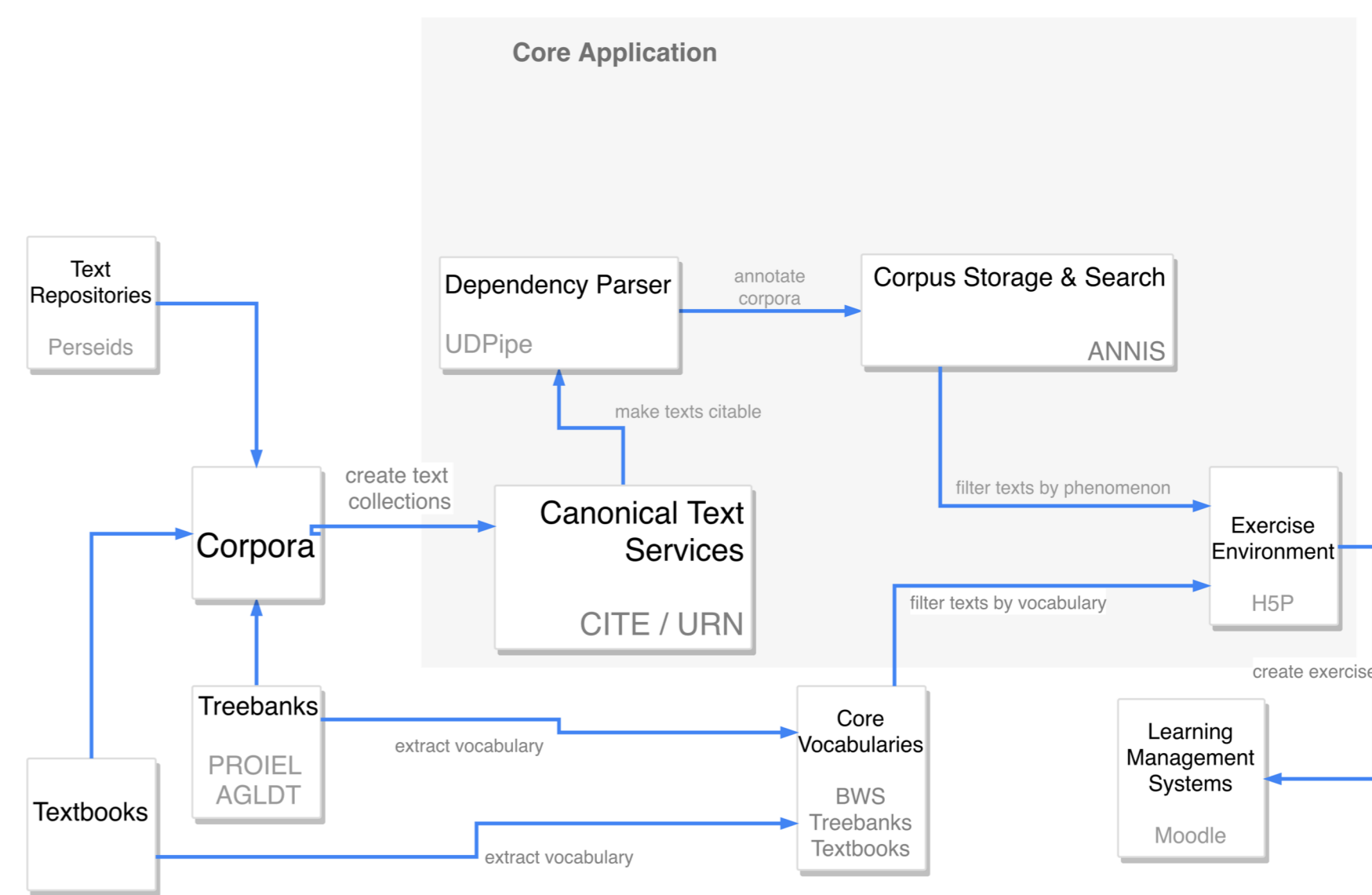


Fig. 1: Existing resources and their processing in the Machina Callida pipeline

Fig. 3: Software Re-Use in the Machina Callida

Machina Callida: Software Re-Use



Try it yourself!

Machina Callida is...

- **Modular:** server interface easily extendable by new functionality
- **Mobile:** hybrid web application that supports mobile learning
- **Efficient:** graph databases for corpus storage and search
- **Sustainable:** installation as container for increased stability and comfort
- **Collaborative:** open source code using common languages and frameworks
- **Shareable:** free license (MIT)

Fig. 2: Example exercise generated by the Machina Callida

Matching: Assign the matching elements to each other!

| | | |
|------------|-------|--------------|
| finis | _____ | Romani |
| itineribus | _____ | Compluribus |
| locis | _____ | superioribus |
| periculo | _____ | magno |
| populi | _____ | suos |
| proeliis | _____ | magnis |
| rebus | _____ | novis |

[Download](#) [Embed](#)

Results

Machina Callida already...

- generates interactive **cloze** and **matching** exercises
- provides exercises with increasing degrees of **complexity**
- exports exercises to various **platforms** (mobile, learning management systems, print)
- applies **language filters** using pre-defined or custom vocabularies
- **annotates** Latin texts dynamically with a range of phenomena: morphology, lemma, part of speech, syntactic dependencies
- **visualizes** syntactic information
- extends its **citation** model to custom corpora, e.g. textbooks or treebanks

Prospects for future research

- Treebanks: work pretty well for creating exercises, but only very **limited amount of texts**
- **Challenges:**
 - Handling of words with **identical spelling** in exercise solutions
 - **No direct feedback** or hinting (yet) in the Machina Callida
 - **No hardware** for manual testing (tablets, mobile phones etc.)
 - **Low accuracy** in existing dependency parsers
 - **User Interface** not (yet) well-designed

References

Bamman, D., & Crane, G. (2011). The Ancient Greek and Latin Dependency Treebanks [AGLDT]. In Language Technology for Cultural Heritage (pp. 79–98). Springer. Retrieved from https://perseusdl.github.io/treebank_data/

Blackwell, C., & Smith, N. (Eds.). (2015). The Canonical Text Services URN Specification, Version 2.0.rc.1 [CITE / URN]. Retrieved from http://cite-architecture.github.io/ctsum_spec/

Clérice, T. (n.d.). Perseids Canonical Text Services. Perseus Digital Library. Retrieved from <https://cts.perseids.org/read/latinLit>

H5P. (n.d.). Joubel. Retrieved from <https://h5p.org/about-the-project>

Haug, D. T., & Jøhndal, M. (2008). Creating a parallel treebank of the old Indo-European Bible translations [PROIEL]. In Proceedings of the Second Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2008) (pp. 27–34). Retrieved from <http://www.hf.uio.no/ifikk/english/research/projects/proiel/Activities/proiel/publications/marrakech.pdf>

Krause, T. (2019). ANNIS: A graph-based query system for deeply annotated text corpora. <https://doi.org/10.5281/zenodo.2598164>

Moodle. (n.d.). (Version 3.6). Retrieved from <https://moodle.org/>

Straka, M., & Straková, J. (n.d.). UDPipe (Version 2.0 CONLL17). A LINDAT/CLARIN project. Retrieved from <http://hdl.handle.net/11234/1-1990>

Utz, C. (2000). Mutter Latein und unsere Schüler — Überlegungen zu Umfang und Aufbau des Wortschatzes [BWS]. Antike Literatur—Mensch, Sprache, Welt. Dialog Schule Und Wissenschaft, 34, 146–172. Retrieved from https://www.ccbuchner.de/_files_media/mediathek/downloads/267.pdf

Contact

Project Website: <https://hu.berlin/callidus>
 Source Code: <https://scm.cms.hu-berlin.de/callidus>
 Public Software Installation: <https://korpling.org/mc/> (v0.6.5)

DFG Project number: 316618374
 Running time: 2017-2020
 E-mail: konstantin.schulz@hu-berlin.de

Address:
 Konstantin Schulz
 Humboldt-Universität zu Berlin
 Sprach- und Literaturwissenschaftliche Fakultät
 Institut für deutsche Sprache und Linguistik
 Unter den Linden 6
 10099 Berlin