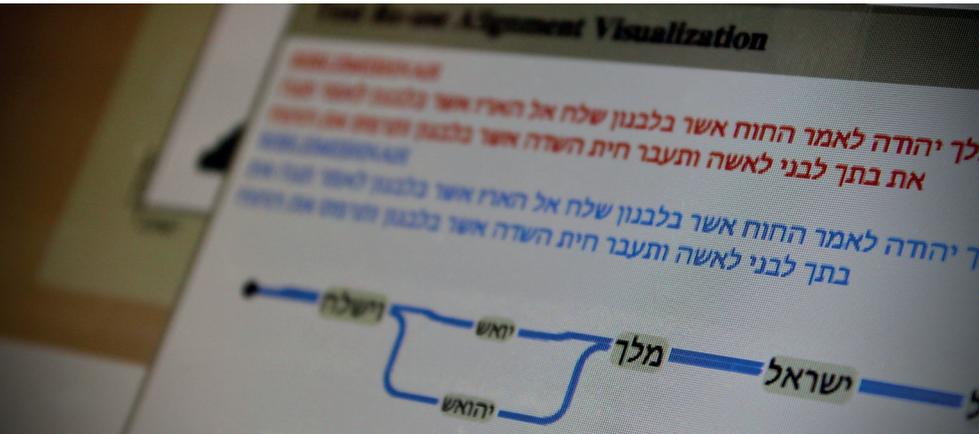


## Zitate und Paraphrasen einfach in großen Datenbeständen wiederfinden Finding quotations and paraphrases in huge amounts of data



Variant Graph einer Text-Reuse-Alignment  
Variant graph of a text reuse alignment

Sokrates ist einer der bekanntesten Philosophen, obwohl er niemals eigene Texte verfasst hat. Lediglich dank seiner Schüler sind seine Lehren verschriftlicht worden und haben sich „viral“ verbreitet.

TRACER ist eine Software zur Erkennung von derartigem Text Reuse in Form von Zitaten, Paraphrasen oder auch Allusionen. Die Architektur basiert auf einem 6-Ebenen-Modell. Jede Ebene kann als ein Arbeitsschritt verstanden werden. Die Ergebnisse eines Arbeitsschrittes werden jeweils zwischengespeichert, so dass diese für spätere Analysen wiederverwendet

werden können. Insgesamt stehen mehrere hundert Algorithmen zur Verfügung, welche zu über eine Million Möglichkeiten miteinander kombiniert werden können.

TRACER ist sprachenunabhängig und auf über zehn Sprachen bereits produktiv im Einsatz. Es wurde unter Berücksichtigung historischer Sprachen, wie Altgriechisch, Latein und Koptisch entwickelt, die sich teilweise über mehr als 2000 Jahre erstrecken und somit eine sprachliche Vielfalt von verschiedenen Schreibweisen oder auch semantischer und morphologischer Sprachevolution abdecken, die es in

modernen Sprachen durch Sprachnormierungen nicht annähernd gibt.

TRACER wird seit 2010 stetig weiterentwickelt und ist dank der Lernkurve auf historischen Sprachen, robust für moderne Anwendungen wie der Erkennung von Plagiarismus oder auch dem Single Sourcing von unternehmensweiten Dokumenten.

TRACER is a software framework for automatic text reuse detection (e.g., quotations or paraphrase). It is based on a six-step model and comprises hundreds of algorithms, which can be combined into over one million different permutations.

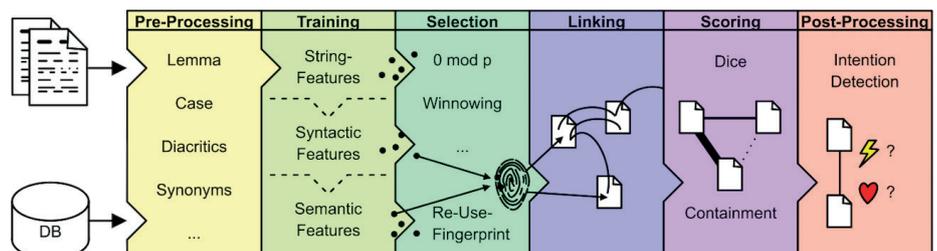
TRACER is language-independent and in productive use in a dozen different languages. Developed for historical languages such as Ancient Greek, Latin, Classical Arabic and Coptic, some of which span millennia and thus embody a complex linguistic evolution, TRACER is also a robust solution for modern applications, including the detection of plagiarism or the single sourcing of company-wide documents.

### Kontakt

Georg-August-Universität Göttingen  
Institut für Informatik  
eTRAP Research Group

Dr. Marco Büchler

contact@etrap.eu  
http://www.etrap.eu



6-Ebenen-Modell der TRACER Architektur  
6 level model of the TRACER architecture