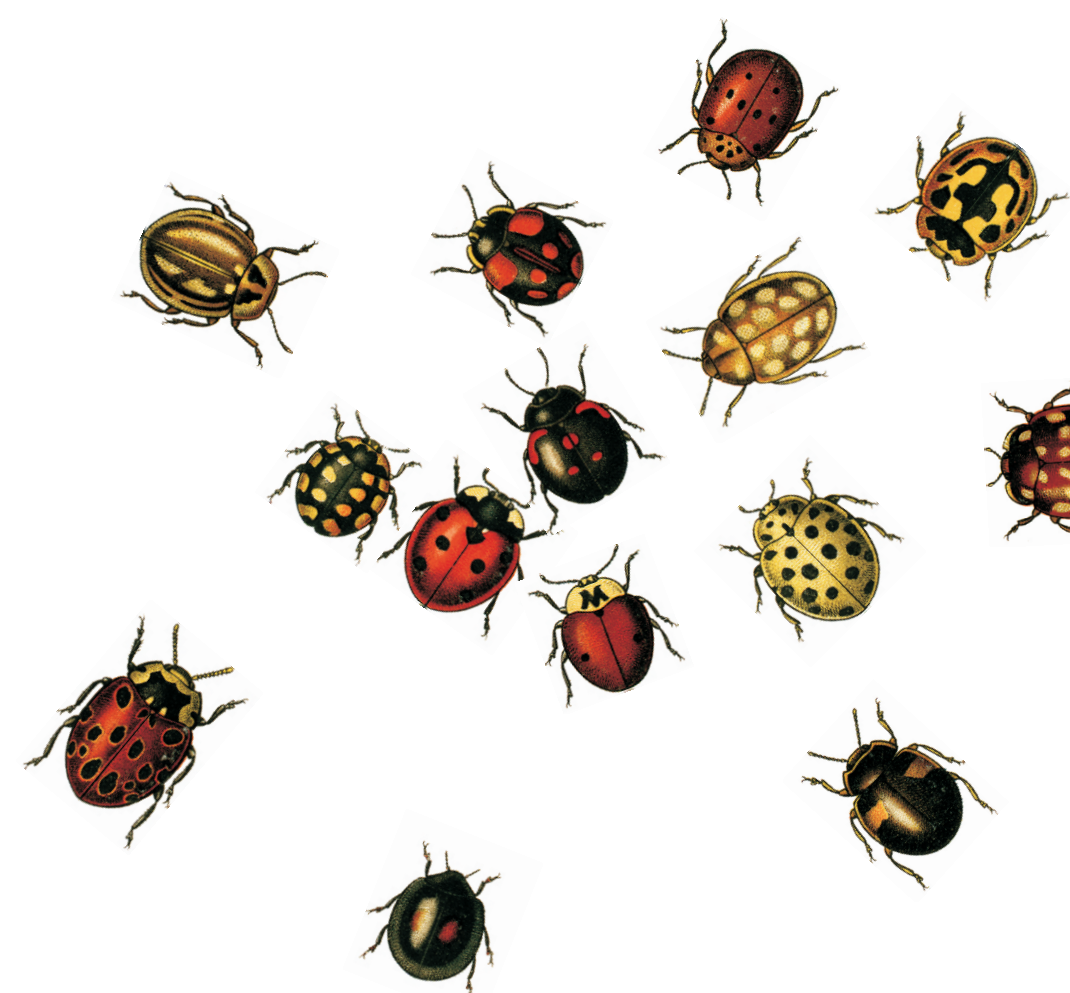


Wechselwirkungen zwischen linguistischen und bioinformatischen Verfahren, Methoden und Algorithmen:

Modellierung und Abbildung von Varianz in Sprache und Genomen (Verbundprojekt)



keffer **Käwweer**
kevere keiver **Kewer**
Käfer kafer
chevar **chevaro** kiäwe
keifer

Untersuchung der **Gemeinsamkeiten** von Genomcode und sprachlichem Code

Kompetenzzentrum für elektronische Erschließungs- und Publikationsverfahren in den Geisteswissenschaften, Universität Trier

Hier liegen digitalisierte Wörterbücher vor, die synchrone wie diachrone Varianz widerspiegeln und damit eine geeignete Datengrundlage für deren Untersuchung liefern. In einer Metalemmaliste werden etymologisch verwandte Lemmata über einen neuhochdeutschen Standard miteinander in Bezug gesetzt, was die methodisch-interdisziplinäre Kompatibilität ermöglicht.

Lehrstuhl für Informatik I - Informationsstrukturen und wissensbasierte Systeme / Institut für deutsche Philologie, Computerphilologie, Universität Würzburg

Ausgehend von effizienten Speicherstrukturen wie XML/TEI, entwickelt der Teilbereich der Informatik mit deklarativen Methoden Techniken zum Annotieren von Wörterbüchern, zur Morphemzerlegung sowie automatische Zuordnungsverfahren für die Metalemmaliste.

Als Ausgangsbasis für eine ausreichende Zerlegung und Zuordnung werden die im Vorfeld annotierten historischen Wörterbuchdaten verwendet und durch kontextsensitive Grammatiken analysiert. Durch automatische Verfahren sollen Gesetzmäßigkeiten in der Evolution der Sprache gefunden und für weitere Verarbeitungsschritte genutzt werden.

Gruppe für Computergestützte Evolutionsbiologie, Abteilung für Bioinformatik, Biozentrum Universität Würzburg

Der Fokus der Gruppe ist die Frage 'Wie entsteht Variabilität?'. Wir wollen verstehen, wie Proteine evolvieren und wie sich dies auf Proteininteraktionen auswirkt. Im 'Wechselwirkungsprojekt' wenden wir bioinformatische Verfahren auf Sprachdaten an. Wir untersuchen, ob die Mechanismen, die die Varianz erzeugen, analog zu biologischen Mechanismen sind. Wir konnten bereits Analogien zwischen Morphemen, Untereinheiten von Wörtern, und Domänen, Bausteinen von Proteinen, identifizieren.

Institut für Deutsche Sprache (IDS), Mannheim

Am IDS wird eine Basislemmaliste des Standardneuhochdeutschen als zentraler Bestandteil der Metalemmaliste geschaffen. Ein methodischer Schwerpunkt ist eine Neuorientierung bei der Erstellung von Wortlisten. Geplant ist die Entwicklung eines Multiagentensystems zur Wortlisten-Generierung, das dem komplexen Charakter von Sprache, die sich ähnlich wie biologische Systeme nicht durch singuläre Entscheidungsträger, sondern durch das Zusammenspiel verschiedenartiger 'Agenten' konstituiert, gerecht wird.