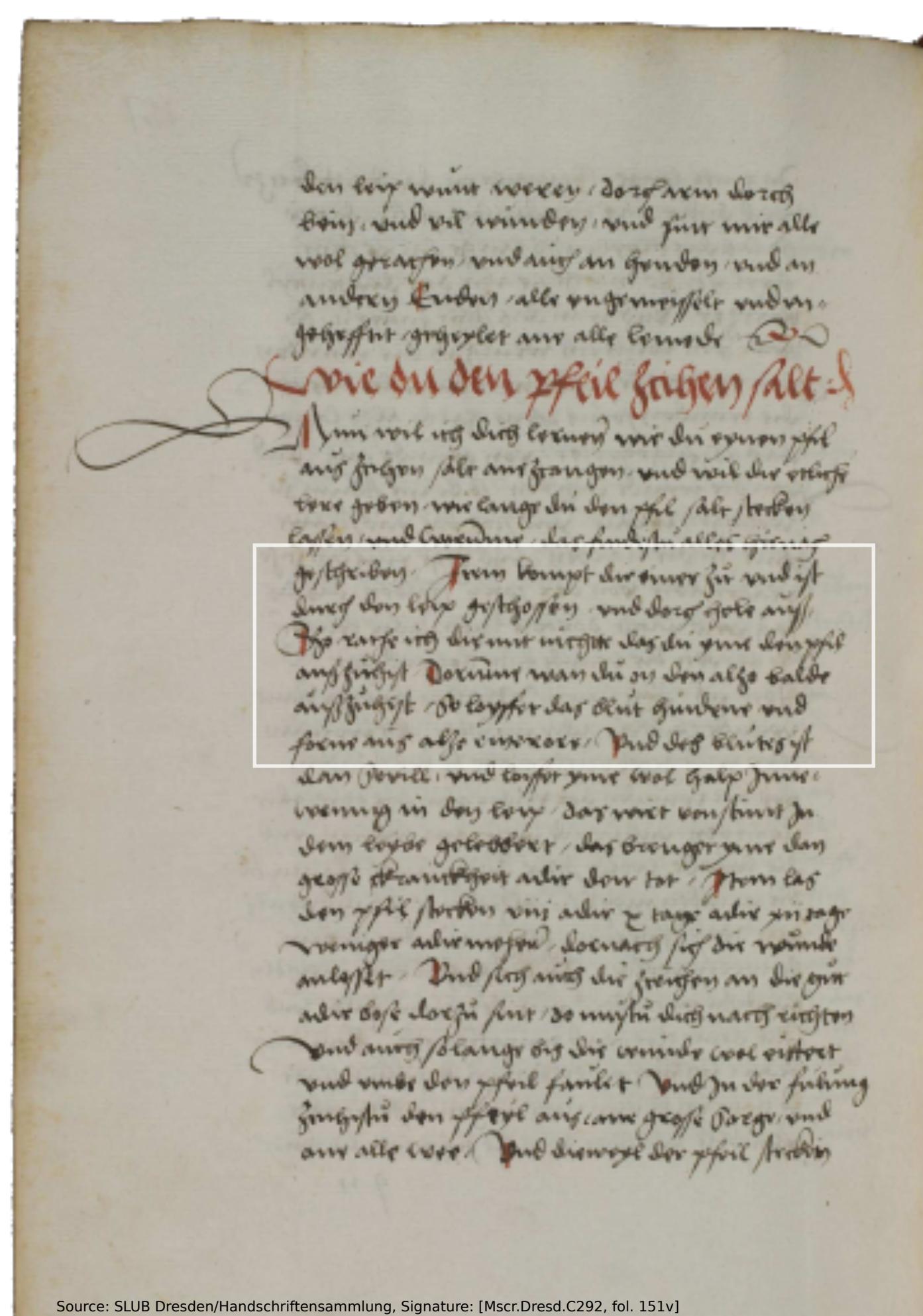


User-friendly lemmatization and morphological annotation of Early New High German manuscripts

SaDA - Semiautomatic difference analysis of complex text variants

SaDA is a BMBF funded project in which germanists, romanists and computer scientists work together aiming at the development of software tools for comparing multiple records or variants of a text. Results are to be presented according to philological scientists needs aiming at critical editions and on-line editions with synopsis and critical apparatus. <http://informatik.uni-halle.de/SaDA>



Source: SLUB Dresden/Handschriftensammlung, Signature: [Mscr.Dresd.C292, fol. 151v]

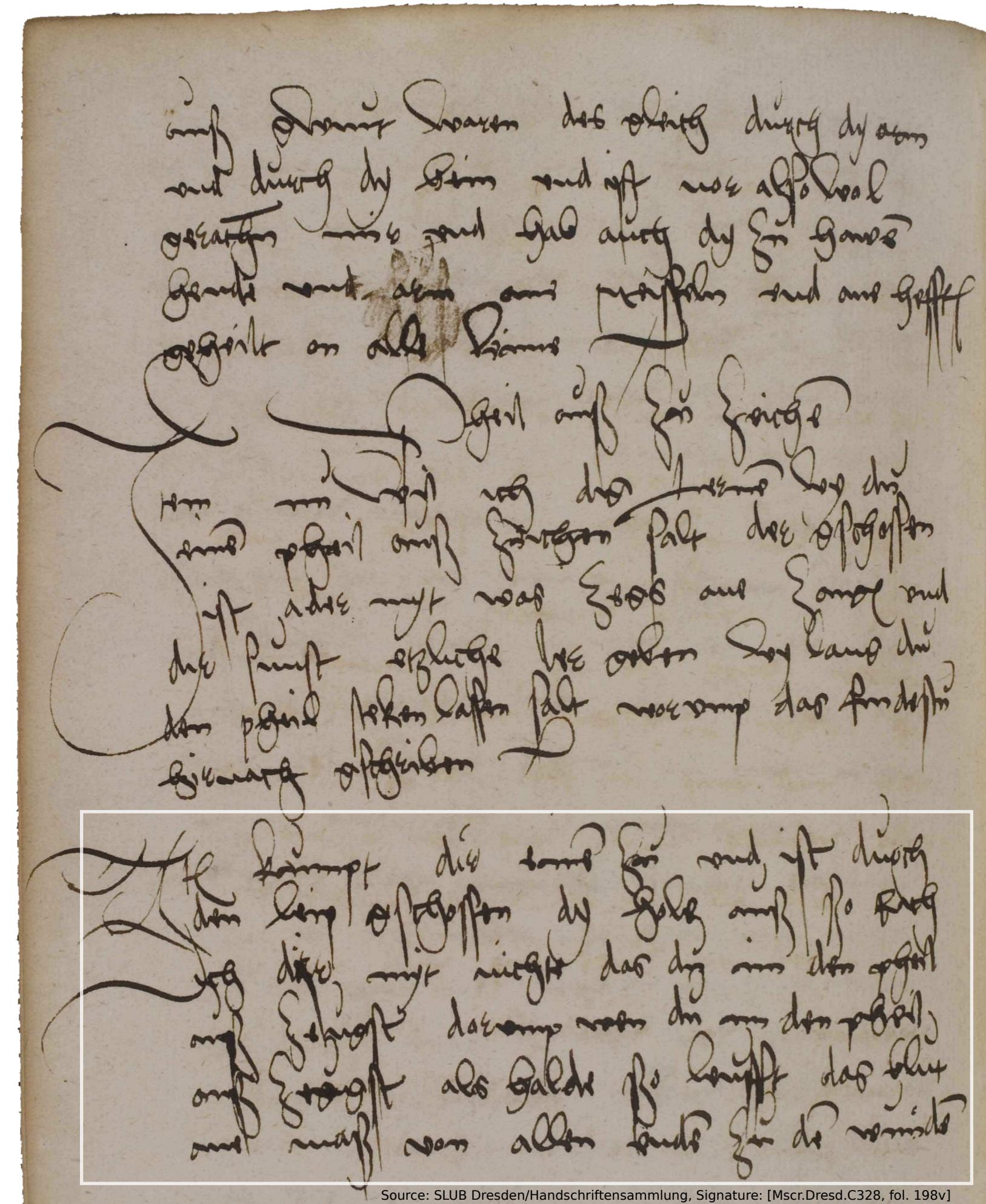
- 6 D2-151v11 geſchriven/(.)§ *f* <j*>tem(.) kompt*1a dir einer zu*§*2a vnd i\$
- 7 D2-151v12 durch den leip geſchoſſen/§ vnd durch hole au*§(.§)
- 8 D2-151v13 *<S*>zo rathe ich dir mit nichete(.)§ das du* yme den pfil
- 9 D2-151v14 au*§*2b zu*hiſt*1b(.)§ *<D*>oru*§m* e(.) wan du* <e>n den alzo balde
- 10 D2-151v15 au*§*2a zu*hiſt*1a/(.)§ So loyſſet*1b das blu*t hindene vnd
- 11 D2-151v16 forme au*§*2b also eine|rore/*<V*>nd des blu*tes i\$

The "Wundarznei" of Heinrich von Pfalzpaint

Heinrich von Pfalzpaint is known as the forefather of plastic surgery, and one of the most famous surgeons of the late middle ages. Of his famous medical compendium, the "Wundarznei" of 1460, eleven manuscript records by different copyists of its time are known. There are ten records available, each having up to 247 chapters and all together having about 700000 wordforms.

Transcription

LAKomp supports a transcription notation which has been developed by germanists from Bonn, Bochum and Halle for the encoding of old German manuscripts in a readable and concise manner.



Source: SLUB Dresden/Handschriftensammlung, Signature: [Mscr.Dresd.C328, fol. 198v]

- 8 D3-198v13 *f*{I*}t\(.) ku*mpt*1a dir eine' zu*§*2a vnd i\$ du*rch
- 9 D3-198v14 den leip geſchoſſen/§ dy hole au*§(.§) & Rath
- 10 D3-198v15 ich d<I>r my.t nichete(.)§ das du* im den pheil
- 11 D3-198v16 au*§*2b zeu*§t*1b(.)§ darvmp(.) wen du* im den pheil
- 12 D3-198v17 au*§*2a zeu*§t*1a +K zeug*§t*1a: eu aus eg gebess. @K als#balde(.)§ &o...
- 13 D3-198v18 aue maß von allen ende\ zu* de' wu*ndle(-.)§

LAKomp: Lemmatization, Annotation, Comparison

The screenshot shows a list of words from a manuscript, each with a small box containing its transcription. To the right, a detailed annotation dialog is open for the word 'ziehen'. It includes fields for 'Lemma Lookup in Grimm' (set to 'aus~ ziehen'), 'POS-tag' (VVIMP: sg.), 'Numerus' (sg.), 'Lemma Note', and 'Note'. A checkbox 'Annotation still unconfirmed' is checked. Buttons for 'Save Annotation' and 'Cancel' are at the bottom right.

Spelling Tolerance

Generated spelling variants for "zeugft":
zeugth, zäugth, ziegth, zeugfh, zäugfh, ziegfh,
zeugszth, zäugszth, ziegzszt, zeugsth, zäugsth,
ziegsth, zeugt, zäugt, ziegts, zäugft, ziegft, zeugszt,
zäugszt, ziegzszt, zeugst, zäugst, ziegst.

LAKomp reduces the manual annotation effort significantly by supporting the user with intuitive dialogs and annotation suggestions. During the annotation process the tool learns from the user's input. Thus the suggestions are improved during the annotation process which helps the user to focus on his annotation decisions.

Detailed Comparison

Using lemmata and morphological data a detailed comparison of witnesses can be generated automatically.

The screenshot shows a comparison of lines from Heinrich von Pfalzpaint's 'Wundarznei'. The top line is 'Meister, Schüler'. Below it, several lines of text are compared, with annotations in red and blue. A legend at the bottom explains the color coding for different witnesses and annotations.

Zeile 18: Meister, Schüler

H Dar vmb/, 1 wen 23 du in 4 5 6 7 alfo 8 bald 9 aub zewheft, 10 11 12
Bu dorumb, wen du yn zo baldes als cogeft,
D3 darvmp, wen du im den pheil aub zeogft alsbalde,
E dar dmb, wan dd in so bald ztcecheft,
Br darvmp, wen du on alfo baldes aub cogeft,
B8 darvmb, wan du ymē den pfi alzo balde uz czuhift,
D2 Dorame, wan du en den alzo balde aub zahift/
P dorumb, wan du im denn pfeil alfo palde als czedgeft.
B9 darvn, wan du jn den pfi vs zugift alfo bald,
St

¹ Dar vmb/, 1 <fehlt> St dar ümb, E, P darvmp, B8 darvmp, D3 darvñ, B9 dorumb, Bu Dorame, D2
² wen I <fehlt> E, B8, D2, P B9, St ³ <fehlt> wan E, B8, D2, P B9
⁴ du In I <fehlt> St dā en D2 dā im D3, P du jm B9 du ymē B8 dā yn Bu da ör Br
⁵ <fehlt> I pfeil P pfi B8, B9 pfeil D3 ⁶ alfo I <fehlt> Bu, D3, E, B9, St alzo B8, D2 alfo Br
⁷ bald I <fehlt> D3, B9, St balde Bu, Br, B8, D2 palde P
⁸ auß zewheft/, 1 <fehlt> St adz czeugeft/P adz cogeft Bu auß zeugft D3 aub zähift/D2 aufs zcogeft, Br
⁹ auß zuecheft/, E uſz czuhift, B8 vs zugift B9 ¹¹ <fehlt> alfo bald, B9 ¹² <fehlt> alsbalde, D3



MARTIN-LUTHER-UNIVERSITÄT
HALLE-WITTENBERG

Institute of Computer Science

André Gießler – andre.giessler@informatik.uni-halle.de
Jörg Ritter – joerg.ritter@informatik.uni-halle.de
Paul Molitor – paul.molitor@informatik.uni-halle.de

Institute of German Studies

Hans-Joachim Solms – hans.joachim.solms@germanistik.uni-halle.de
Sylwia Kösser – sylwia.koesser@germanistik.uni-halle.de
Aletta Leipold – aletta.leipold@germanistik.uni-halle.de



GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung