

## Journal Articles

Anna Shadrova, Anke Lüdeling, Martin Klotz, Rahel Gajaneh Hartz, and **Thomas Krause**. „Step away from the Computer!“. Über die linguistische Datenkategorisierung als Erkenntnisprozess und daraus folgende Herausforderungen bei der Nachnutzung von Annotationen und Annotationstools“. In: *Zeitschrift für germanistische Linguistik* 53.1 (10 April 2025), pages 166–214. DOI: 10.1515/zgl-2025-2005.

Stephan Druskat, Nasir U. Eisty, Robert Chisholm, Neil Chue Hong, Ryan C. Cocking, Myra B. Cohen, Michael Felderer, Lars Grunske, Sarah A. Harris, Wilhelm Hasselbring, **Thomas Krause**, Jan Linxweiler, and Colin C. Venters. „Better Architecture, Better Software, Better Research“. In: *Computing in Science & Engineering* (2025), pages 1–11. DOI: 10.1109/MCSE.2025.3573887.

Anke Lüdeling, Luka Szucsich, Lars Erik Zeige, Aria Adli, Artemis Alexiadou, Malte Belz, Miriam Bouzouita, Oliver Bunk, Malte Dreyer, Markus Egg, Anna Helene Feulner, Jürg Fleischer, Natalia Gagarina, Aron Hirsch, Stefanie Jannedy, Pia Knoeferle, **Thomas Krause**, Silvia Kutscher, Mingya Liu, Beate Lütke, Antonio Priemer, Katja Maquate, Laura Merino Hernández, Roland Meyer, Christine Mooshammer, Stefan Müller, Uli Sauerland, Antje Saueremann, Viola Schmitt, Nicole Schumacher, Dina Serova, Stephanie Solt, Jozina Vander Klok, Elisabeth Verhoeven, Richard Waltereit, Melanie Weirich, and Heike Wiese. „Register: Language Users’ Knowledge of Situational-Functional Variation“. In: *Register Aspects of Language in Situation* 3.1 (2024), pages 1–53. DOI: 10.18452/29131.

Stephan Druskat, **Thomas Krause**, Clara Lachenmaier, and Bastian Bunzeck. „Hexatomic: An extensible, OS-independent platform for deep multi-layer linguistic annotation of corpora“. In: *Journal of Open Source Software* 8.86 (2023), page 4825. DOI: 10.21105/joss.04825.

Anke Lüdeling, Artemis Alexiadou, Aria Adli, Karin Donhauser, Malte Dreyer, Markus Egg, Anna Helene Feulner, Natalia Gagarina, Wolfgang Hock, Stefanie Jannedy, Frank Kammerzell, Pia Knoeferle, **Thomas Krause**, Manfred Krifka, Silvia Kutscher, Beate Lütke, Thomas McFadden, Roland Meyer, Christine Mooshammer, Stefan Müller, Katja Maquate, Muriel Norde, Uli Sauerland, Stephanie Solt, Luka Szucsich, Elisabeth Verhoeven, Richard Waltereit, Anne Wolfgruber, and Lars Erik Zeige. „Register: Language Users’ Knowledge of Situational-Functional Variation“. In: *Register Aspects of Language in Situation* 1.1 (29 July 2022), pages 1–58. DOI: 10.18452/24901.

Carolin Odebrecht, Malte Belz, Amir Zeldes, Anke Lüdeling, and **Thomas Krause**. „RIDGES Herbology: designing a diachronic multi-layer corpus“. In: *Language Resources and Evaluation* 51.3 (September 2017), pages 695–725. DOI: 10.1007/s10579-016-9374-3.

**Thomas Krause**, Ulf Leser, and Anke Lüdeling. „graphANNIS: A Fast Query Engine for Deeply Annotated Linguistic Corpora“. In: *JLCL* 31.1 (2016), pages Iii–25. DOI: 10.21248/jlcl.31.2016.199.

**Thomas Krause** and Amir Zeldes. „ANNIS3: A new architecture for generic corpus query and visualization“. In: *Digital Scholarship in the Humanities* 31.1 (2016), pages 118–139. DOI: 10.1093/llc/fqu057.

## Book Sections

Esra Uyanik, Sven Oliver Müller, Anke Lüdeling, and **Thomas Krause**. „Differenzierung und Standardisierung: Zur Entwicklung von Registern“. In: *Historische Fach- und Wissenschaftstexte kontrastiv*. De Gruyter, 6 May 2025. pages 193–214. URL: <https://www.degruyterbrill.com/document/doi/10.1515/9783111328867-010/html> (visited on 2025-04-23).

Martin Klotz and **Thomas Krause**. „Eine Ressource durch Abwägung: Wie bestehende generische Softwarelösungen ‚über sich hinauswachsen‘ und ein Parallelkorpus ermöglichen“. In:

*Mittelniederdeutsch zwischen Korpuslinguistik und Literaturwissenschaft*. Edited by Marco Coniglio, Anabel Recker, and Heike Sahn. Göttingen: Göttingen University Press, 2024. pages 123–138. DOI: 10.17875/gup2024-2620.

## Conference Papers

**Thomas Krause**, Ulf Leser, Anke Lüdeling, and Stephan Druskat. “Designing a Re-Usable and Embeddable Corpus Search Library”. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*. 6th Workshop on the Challenges in the Management of Large Corpora. Edited by Hanno Biber, Evelyn Breiteneder, Simon Clematide, and Andreas Witt. Miyazaki, Japan: European Language Resources Association (ELRA), 12 July 2018. URL: [http://lrec-conf.org/workshops/lrec2018/W17/pdf/12\\_W17.pdf](http://lrec-conf.org/workshops/lrec2018/W17/pdf/12_W17.pdf).

Carolin Odebrecht, Malte Dreyer, Anke Lüdeling, and **Thomas Krause**. “Zwei grundlegende Fragen der digitalen Nachhaltigkeit: Wie können wir die heterogenen Forschungsfragen und die Community bei der Verfügbarmachung von Forschungsdaten miteinbeziehen?” In: *Digitale Nachhaltigkeit. DHD 2017*. Bern, February 2017. URL: [http://www.dhd2017.ch/wp-content/uploads/2017/01/Abstractband\\_def\\_24.1.17-1.pdf](http://www.dhd2017.ch/wp-content/uploads/2017/01/Abstractband_def_24.1.17-1.pdf).

Stephan Druskat, Volker Gast, **Thomas Krause**, and Florian Zipser. “corpus-tools.org: An Interoperable Generic Software Tool Set for Multi-layer Linguistic Corpora”. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*. Edited by Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis. Portorož, Slovenia: European Language Resources Association (ELRA), May 2016. URL: [http://www.lrec-conf.org/proceedings/lrec2016/pdf/918\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2016/pdf/918_Paper.pdf).

**Thomas Krause**, Julia Ritz, Amir Zeldes, and Florian Zipser. “Topological Fields, Constituents and Coreference: A New Multi-layer Architecture for TüBa-D/Z.” In: *Proceedings of the Conference of the German Society for Computational Linguistics and Language Technology (GSCL) 2011*. GSCL 2011. Hamburg, 2011. pages 259–262. URL: <http://exmaralda.org/g scl2011/downloads/AZM96.pdf>.

Heinrich Mellmann, Yuan Xu, **Thomas Krause**, and Florian Holzhauser. “NaoTH Software Architecture for an Autonomous Agent”. In: *Proceedings of the International Workshop on Standards and Common Platforms for Robotics (SCPR 2010)*. Darmstadt, November 2010. pages 316–327. URL: <http://www.naoTEAMhumboldt.de/wp-content/papercite-data/pdf/scpr-mellmannxuetal-10.pdf>.

## Talks and Posters

Martin Klotz, **Thomas Krause**, Anke Lüdeling, Carolin Odebrecht, and Malte Dreyer. “Large language models and the process of corpus linguistic research”. Corpus Linguistics & AI (Rauischholzhausen). 8 May 2026.

Stephan Druskat and **Thomas Krause**. “The four elements of achieving research software sustainability for long tail projects”. deRSE23 - Conference for Research Software Engineering in Germany (Paderborn). 19 February 2023. DOI: 10.5281/zenodo.7654777.

Anke Lüdeling, Roland Schäfer, Elizabeth Pankratz, **Thomas Krause**, Felix Bildhauer, and Felix Golcher. “What is a good corpus?” (Berlin). March 2021. URL: <https://rolandschaefer.net/wp-content/uploads/2021/06/GoodCorpusSFB1412Handout.pdf>.

**Thomas Krause** and Stephan Druskat. “Die Hard 1.1024.0: backward compatibility of a search engine with persistent IDs”. deRSE19 - Conference for Research Software Engineers in Germany (Potsdam). 4 June 2019. DOI: 10.5446/42501.

Stephan Druskat, **Thomas Krause**, and Carolin Odebrecht. "Utilising ANNIS for search and analysis of historical data". Reuse or New Development: sustainability of resources and tools for multi-faceted historical data and languages. FORGE 2016 (Hamburg). September 2016. DOI: 10.5281/zenodo.160368.

Stephan Druskat, **Thomas Krause**, Carolin Odebrecht, and Florian Zipser. "A workflow for creating, analysing, and storing multi-layer corpora: Pepper, Atomic, ANNIS and LAUDATIO". In: *DGfS-CL Poster Session. 38. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. 2016. DOI: 10.5281/zenodo.47345.

**Thomas Krause**, Anke Lüdeling, Carolin Odebrecht, and Dennis Zielke. "LAUDATIO. Ein flexibles Repositorium für historische Textdaten". FORGE 2015 (Hamburg). September 2015. URL: <https://www.fdm.uni-hamburg.de/ueber-uns/a-nachrichten/aktivitaeten/forge15/programm/folien/forge2015abstracts.pdf>.

Florian Zipser, **Thomas Krause**, Anke Lüdeling, Arne Neumann, Manfred Stede, and Amir Zeldes. "ANNIS, SaltNPepper & PAULA: A multilayer corpus infrastructure". In: *Final Conference of the SFB 632 Information Structure: Advances in Information Structure Research 2003 - 2015*. Berlin, May 2015. DOI: 10.5281/zenodo.20713.

Florian Zipser, André Röhrig, Anke Lüdeling, Martin Klotz, **Thomas Krause**, Stephan Druskat, and Vivian Voigt. "SaltNPepper, ANNIS & Atomic: Eine Infrastruktur für Mehrebenenkorpora". Poster. (Berlin). May 2015. URL: <http://dx.doi.org/10.5281/zenodo.20714>.

Florian Zipser, **Thomas Krause**, and Arne Neumann. "Interactive analysis of multi-layer linguistic corpora with ANNIS". In: *DGfS-CL Poster Session. 37. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. April 2015. DOI: 10.5281/zenodo.16663.

Carolin Odebrecht, **Thomas Krause**, and Anke Lüdeling. "Austausch von historischen Texten verschiedener Sprachen über das LAUDATIO-Repository". In: *DGfS-CL Poster Session. 37. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. Universität Leipzig, March 2015. URL: <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiterinnen/carolin/dgfs2015>.

**Thomas Krause**, Carolin Odebrecht, and Dennis Zielke. "Langfristiger Zugang und wissenschaftliche Nutzung historischer Korpora mit der korpuslinguistischen Infrastruktur LAUDATIO". In: *Open-Access-Strategie für Berlin*. Berlin: Freie Universität zu Berlin, October 2014. URL: [https://www.fu-berlin.de/sites/open\\_access/weiteres/Veranstaltungen/oa\\_berlin/poster/LAUDATIO\\_Krause-Thomas-Odebrecht-Carolin-Zielke-Dennis\\_HU-Berlin.pdf](https://www.fu-berlin.de/sites/open_access/weiteres/Veranstaltungen/oa_berlin/poster/LAUDATIO_Krause-Thomas-Odebrecht-Carolin-Zielke-Dennis_HU-Berlin.pdf).

Carolin Odebrecht, Dennis Zielke, **Thomas Krause**, Benjamin Weißenfels, Malte Belz, Tino Schernikau, and Vivian Voigt. "Wissenschaftliche Nutzung der korpuslinguistischen Infrastruktur LAUDATIO". In: *DGfS-CL Poster Session. 36. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. Marburg, March 2014. URL: [https://web.archive.org/web/20170326230820/http://paralleltxt.info/dgfs2014/abstracts/Odebrecht\\_Zielke\\_Krause\\_Weissenfels\\_Belz\\_Schernikau\\_Voigt.pdf](https://web.archive.org/web/20170326230820/http://paralleltxt.info/dgfs2014/abstracts/Odebrecht_Zielke_Krause_Weissenfels_Belz_Schernikau_Voigt.pdf).

**Thomas Krause**, Anke Lüdeling, Carolin Odebrecht, Laurent Romary, Peter Schirmbacher, and Dennis Zielke. "LAUDATIO-Repository: Accessing a heterogeneous field of linguistic corpora with the help of an open access repository". In: *Digital Humanities 2014 Conference. Poster Session*. 2014. URL: <https://dh-abstracts.library.virginia.edu/works/1981>.

**Thomas Krause**, Benjamin Weißenfels, Amir Zeldes, and Florian Zipser. “ANNIS3: Towards Generic Corpus Search and Visualization”. In: *DGfS-CL Poster Session. 36. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. Marburg, 2014. DOI: 10.5281/zenodo.15639.

**Thomas Krause**, Carolin Odebrecht, Amir Zeldes, and Florian Zipser. “Unary TEI Elements and the Token Based Corpus”. (Sapienza Università di Roma). 1 October 2013. URL: [https://tei-c.org/Vault/Meetings/2013/wp-content/uploads/2013/09/Krause\\_et\\_al.pdf](https://tei-c.org/Vault/Meetings/2013/wp-content/uploads/2013/09/Krause_et_al.pdf).

Carolin Odebrecht and **Thomas Krause**. “Metadata in an Infrastructure for Historical Corpora.” Kolloquium. SFB 732 Incremental Specification in Context (Stuttgart). 20 June 2013. URL: [https://www.sfb732.uni-stuttgart.de/documents/files/abstract\\_odebrechtkrause.pdf](https://www.sfb732.uni-stuttgart.de/documents/files/abstract_odebrechtkrause.pdf).

**Thomas Krause**, Carolin Odebrecht, and Dennis Zielke. “Wie kann der Zugriff, die Wiederverwendung und langfristige Speicherung von linguistischen Korpora realisiert werden?” In: *DGfS-CL Poster Session. 35. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. Universität Potsdam, 2013. URL: <https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiter-innen/carolin/dgfs2013-krause-odebrecht-zielke/>.

**Thomas Krause**, Anke Lüdeling, Carolin Odebrecht, and Amir Zeldes. “Multiple Tokenization in a Diachronic Corpus”. Exploring Ancient Languages through Corpora Conference (EALC) (Universitetet i Oslo). 2012. URL: [https://web.archive.org/web/20201030085904/https://www.hf.uio.no/ifikk/english/research/projects/proiel/ealc/abstracts/Krause\\_et\\_al.pdf](https://web.archive.org/web/20201030085904/https://www.hf.uio.no/ifikk/english/research/projects/proiel/ealc/abstracts/Krause_et_al.pdf).

Marc Reznicek, Cedric Krummes, Hagen Hirschmann, Anke Lüdeling, Astrid Ensslin, Jia Wei Chan, Amir Zeldes Thomas Krause, and Florian Zipser. “Dass wenn man etwas will, muss man dafür arbeiten“- Zielhypothesen im Lernerkorpus Falko”. In: *DGfS-CL Poster Session. 32. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. Berlin, 2010.

Christian Chiarcos, **Thomas Krause**, Anke Lüdeling, Julia Ritz, Viktor Rosenfeld, Manfred Stede, Amir Zeldes, and Florian Zipser. “Search and Visualization of Richly Annotated Corpora with ANNIS2”. In: *DGfS-CL Poster Session. 31. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft (DGfS)*. Osnabrück, 2009. URL: [https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiter-innen/ehemalige/amir/pdf/DGfS09\\_annis\\_poster](https://www.linguistik.hu-berlin.de/de/institut/professuren/korpuslinguistik/mitarbeiter-innen/ehemalige/amir/pdf/DGfS09_annis_poster).

## Theses

**Thomas Krause**. “ANNIS: A graph-based query system for deeply annotated text corpora”. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2019. DOI: 10.18452/19659.

**Thomas Krause**. Erfahrungsbasierte Lernmethoden zur visuellen Trajektorienvorhersage humanoider Roboter. Master’s Thesis. Humboldt-Universität zu Berlin, 2011. URL: [https://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiter-innen/thomas/publications/verhalten\\_nao.pdf](https://www.linguistik.hu-berlin.de/institut/professuren/korpuslinguistik/mitarbeiter-innen/thomas/publications/verhalten_nao.pdf).

## Software

**Thomas Krause**. *graphANNIS*. version 4.1.1. January 2026. DOI: 10.5281/zenodo.2598164.

**Thomas Krause** and Martin Klotz. *Annatto*. version 0.47.0. January 2026. URL: <https://github.com/korpling/annatto/>.

**Thomas Krause**, Weißenfels Benjamin, Tom Rütte, Irina Glushanok, Martin Klotz, Shuo Zhang, Amir Zeldes, Fabian Bartels, Stephan Druskat, Adriane Boyd, Egon Stemle, Lari Lampen, and Florian Petran. *ANNIS*. version 4.15.1. October 2025. DOI: 10.5281/zenodo.1212548.

Stephan Druskat, **Thomas Krause**, Clara Lachenmaier, and Bastian Bunzeck. *Hexatomic*. version 1.5.0-SNAPSHOT. October 2024. DOI: 10.5281/zenodo.6900689.

## Corpora

Heike Wiese, Artemis Alexiadou, Shanley Allen, Oliver Bunk, Natalia Gagarina, Kateryna Iefremenko, Esther Jahns, Martin Klotz, **Thomas Krause**, Annika Labrenz, Anke Lüdeling, Maria Martynova, Katrin Neuhaus, Tatiana Pashkova, Vicky Rizou, Tracy Rosemarie, Christoph Schroeder, Luka Szucsich, Wintai Tsehaye, Sabine Zerbian, and Yulia Zuban. *RUEG Corpus*. Zenodo, 24 April 2020. DOI: 10.5281/zenodo.3765218.

Anke Lüdeling, Carolin Odebrecht, **Thomas Krause**, Gohar Schnelle, and Catharina Fischer. *RIDGES Herbiology (Version 9.0)*. version 9.0. 31 March 2020. DOI: 10.34644/laudatio-dev-PySSCnMB7CArCQ9CNKFY.

## Reports

Heinrich Mellmann, Benjamin Schlotter, Steffen Kaden, Philipp Strobel, **Thomas Krause**, and Claas-Norman Ritter. *Berlin United - Nao Team Humboldt: Team Report 2016*. Team Report. Humboldt-Universität zu Berlin, Adaptive Systems Group, 2016. URL: <http://www.naoteamhumboldt.de/wp-content/papercite-data/pdf/2016-Naoth-Report.pdf>.

Hans-Dieter Burkhard, Florian Holzhauer, **Thomas Krause**, Heinrich Mellmann, Claas Norman Ritter, Oliver Welter, and Yuan Xu. *NAO-Team Humboldt 2010*. Team Report. Humboldt Universität zu Berlin, 2010. URL: <http://www.naoteamhumboldt.de/wp-content/uploads/2010/02/NaoTH10Description.pdf>.

Hans-Dieter Burkhard, Florian Holzhauer, **Thomas Krause**, Heinrich Mellmann, Claas Norman Ritter, Oliver Welter, and Yuan Xu. *NAO-Team Humboldt 2009*. Team Report. Humboldt Universität zu Berlin, 2009. URL: [http://www.naoteamhumboldt.de/wp-content/uploads/2010/02/NaoTH09Report\\_final.pdf](http://www.naoteamhumboldt.de/wp-content/uploads/2010/02/NaoTH09Report_final.pdf).

Thomas Röfer, Tim Laue, Michael Weber, Hans-Dieter Burkhard, Matthias Jüngel, Daniel Göhring, Jan Hoffmann, Benjamin Altmeyer, **Thomas Krause**, Michael Spranger, Oskar Stryk, Ronnie Brunn, Marc Dassler, Michael Kunz, Tobias Oberlies, Max Risler, Uwe Schwiegelshohn, Matthias Hebbel, Walter Nisticó, Stefan Czarnetzki, Thorsten Kerkhof, Matthias Meyer, Carsten Rohde, Bastian Schmitz, Michael Wachter, Tobias Wegner, and Christine Zarges. *GermanTeam 2005*. Team Report. German Team, 2005. URL: <http://web.archive.org/web/20110415023802/http://www.germanteam.org/GT2005.pdf>.