

## Named Entity Recognition (NER)

Eder, Elisabeth; [elisabeth.eder@aau.at](mailto:elisabeth.eder@aau.at)

*Named Entity Recognition* (NER) bezeichnet die Erkennung von Eigennamen (*named entities*) in Texten sowie auch deren Klassifizierung in verschiedene Entitätstypen. In den meisten Fällen wird *Named Entity Recognition* als *Sequence Tagging*- oder *Sequence Labeling*-Aufgabe aufgefasst, bei der jedem Token einer Sequenz eine bestimmte Kategorie bzw. ein bestimmter Entitätstyp zugewiesen wird (siehe auch *Part-of-Speech-Tagging*). Standardmäßig zählen Personen, Orte und Organisationen zu diesen Entitätstypen. Je nach Anwendungsfall werden aber auch andere Kategorien, wie zum Beispiel kommerzielle Produktnamen, Werktitel oder Fahrzeuge, berücksichtigt. (Jurafsky/Martin 2009, S. 761–768) Im Deutschen können auch Derivative (z. B. ‘österreichisch’) und partielle Entitäten, die nur einen Teil eines Token betreffen (z. B. KONDE-Beitrag), eine Rolle spielen und dementsprechende Unterkategorien bilden. (Benikova et al. 2014) *Named Entity Recognition* ist in vielen *Libraries* für NLP implementiert, z. B. *spaCy* oder *Natural Language Toolkit* (nltk), und auch *WebLicht* bietet Tools dafür an.

### Literatur:

- Akbik, Alan; Blythe, Duncan; Vollgraf, Roland: Contextual String Embeddings for Sequence Labeling. In: Proceedings of the 27th International Conference on Computational Linguistics COLING. Santa Fe, New Mexico, USA: 2018, S. 1638–1649.
- Benikova, Darina; Biemann, Chris; Reznicek, Marc: NoSta-D Named Entity Annotation for German: Guidelines and Dataset. In: Proceedings of 9th International Conference on Language Resources and Evaluation (LREC ’14) LREC. Reykjavik, Iceland: 2014, S. 2524–2531.
- Benikova, Darina; Yimam, Seid Muhie; Santhanam, Prabhakaran; Biemann, Chris: GermaNER: Free Open German Named Entity Recognition Tool. In: Proceedings of the International Conference of the German Society for Computational Linguistics and Language Technology GSCL: 2015.
- Deroncourt, Franck; Lee, Ji Young; Szolovits, Peter: NeuroNER: an easy-to-use program for named-entity recognition based on neural networks. In: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations EMNLP. Copenhagen, Denmark: 2017, S. 97–102.
- Jurafsky, Daniel; Martin, James H.: Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Upper Saddle River, New Jersey, United States of America: 2009.

- Riedl, Martin; Padó, Sebastian: A Named Entity Recognition Shootout for German. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers) ACL. Melbourne, Australia: 2018, S. 120–125.

**Software:**

weblicht, Natural Language Toolkit (nltk), flair, German NER, GermaNER, NeuroNER

**Verweise:**

spaCy, WebLicht, Part-of-Speech-Tagging, NLP, Tagger

**Projekte:**

Natural Language Toolkit

**Themen:**

Annotation und Modellierung, Natural Language Processing

**Zitervorschlag:**

Eder, Elisabeth. 2021. Named Entity Recognition (NER). In: KONDE Weißbuch. Hrsg. v. Helmut W. Klug unter Mitarbeit von Selina Galka und Elisabeth Steiner im HRSM Projekt "Kompetenznetzwerk Digitale Edition". URL: <https://gams.uni-graz.at/o:konde.141>