

Joachim Daiber

San Francisco Bay Area
650-334-7478

✉ daiber.joachim@gmail.com

Applied scientist focusing on natural language processing and multilinguality with track record of shipping solutions at scale.

Employment

- 09/2021 **Co-Founder at Kailua Labs Inc. (San Francisco, California)**
- 01/2018–08/2021 **Siri Knowledge at Apple Inc. (Cupertino, California)**
Staff ML Engineer (09/2020), Tech Lead for Query Understanding (2019-2021)
Tech lead for query understanding for the Knowledge domain in Siri, Safari and Spotlight. I established datasets, metrics, rigorous experiment management, as well as machine learning and language expansion strategies, leading the successful expansion of the domain from one to 24 languages and establishing the domain as the best-performing information-seeking domain in Siri by share of correct answers, agility and infrastructure. Selected news coverage: [USA Today](#), [TechCrunch](#), [Apple ML research](#).

Education

- 10/2017 **PhD in Machine Translation and Natural Language Processing**
PhD thesis: Typologically Robust Statistical Machine Translation
ILLC, University of Amsterdam, The Netherlands.
- 09/2013 **MSc and MA in Natural Language Processing**
Grade: *Summa cum laude*, Erasmus Mundus LCT double degree
Charles University in Prague (MSc) and University of Groningen (MA)
- 09/2011 **BSc in Computer Science, English Philology**
Free University of Berlin, Germany

Awards and Scholarships

- since 2017 **O-1 Extraordinary Ability and EB-1 Outstanding Researcher**
United States of America
- 2013–2015 **Marie Skłodowska-Curie Fellowship**
EXPERT ITN, European Commission
- 2011–2013 **Erasmus Mundus Scholarship**
European Commission
- 2011–2012 **Charles University Merit Scholarship**
Charles University in Prague (prospěchové stipendium)

Languages and Programming

- Languages German (native), English (fluent), Spanish (conversational), Dutch (basic).
- Advanced Python/NumPy, Bash, Scala, Java, \LaTeX , JavaScript, CSS.
- Intermediate to Basic Golang, Perl, Haskell, PHP, C, C++, Objective-C.

Selected Publications (see also [Google Scholar](#))

Multilingual Question Answering

- 2020 **MKQA: A Linguistically Diverse Benchmark for Multilingual Open Domain Question Answering**

Shayne Longpre, Yi Lu, and Joachim Daiber

To appear in Transactions of the Association for Computational Linguistics.

Machine Translation

- 2016 **Universal Reordering via Linguistic Typology**

Joachim Daiber, Miloš Stanojević, and Khalil Sima'an.

In COLING 2016, the 26th International Conference on Computational Linguistics.

- 2016 **Examining the Relationship between Preordering and Word Order Freedom in Machine Translation**

Joachim Daiber, Miloš Stanojevic, Wilker Aziz, Khalil Sima'an.

In Proceedings of the First Conference on Machine Translation (Research Papers).

Entity Linking

- 2016 **Improving Efficiency and Accuracy in Multilingual Entity Extraction**

500+ citations, see also dbpedia-spotlight.org.

Joachim Daiber, Max Jakob, Chris Hokamp, and Pablo N. Mendes

In Proceedings of the 9th International Conference on Semantic Systems.

Morphology in Word Representations

- 2016 **Splitting Compounds by Semantic Analogy**

Joachim Daiber, Lautaro Quiroz, Roger Wechsler, and Stella Frank.

In Proceedings of the Tenth International Conference on Language Resources and Evaluation.

Dependency Parsing

- 2016 **The Denoised Web Treebank: Evaluating Dependency Parsing under Noisy Input Conditions**

Joachim Daiber and Rob van der Goot.

In Proceedings of the Tenth International Conference on Language Resources and Evaluation.