

## Michael Denkowski

---

Fort Worth, TX  
USA

michael.j.denkowski@gmail.com  
<https://www.mjdenkowski.com>

### Employment

#### **Amazon.com, Inc.**

*Machine Learning Scientist*

June 2015 – April 2018

*Senior Machine Learning Scientist*

April 2018 – Present

Focus: Developing and applying AI foundation models.

#### **Safaba Translation Solutions, Inc.**

*Research Engineer*

January 2010 – May 2013

*Research Scientist*

May 2013 – June 2015

Focus: Adapting machine translation to client-specific use cases.

### Education

#### **Carnegie Mellon University**

May 2010 – May 2015

*Ph.D. in Language and Information Technology*

Thesis: “Machine Translation for Human Translators”

Department: Language Technologies Institute

Advisor: Alon Lavie

#### **Carnegie Mellon University**

August 2008 – May 2010

*M.S. in Language Technologies*

Department: Language Technologies Institute

Advisor: Alon Lavie

#### **Texas Christian University**

August 2005 – May 2008

*B.S. in Computer Science*

Department: Computer Science

Advisors: Charles Hannon and Antonio Sanchez

### Teaching

#### **Algorithms for Natural Language Processing**

Course LTI 11-711, Carnegie Mellon University

*Teaching Assistant*

Fall 2011

*Teaching Assistant*

Fall 2012

#### **Guest Lectures**

*Building Great Software with Small Teams*, Texas Christian University (Annual 2019-2024)

*Learning to Translate Human Language*, Texas Christian University (2019)

*Machine Translation and Post-Editing*, Carnegie Mellon University (2014, 2016)

*Phrase-Based Machine Translation*, Carnegie Mellon University (2014)

### Publications

*Citations: 4701, h-index: 21, i10-index: 26 (Source: Google Scholar, February 2025)*

## Ph.D. Thesis

*Machine Translation for Human Translators*, Michael Denkowski. Ph.D. Thesis, Carnegie Mellon University, April 2015. Committee: Alon Lavie, Chris Dyer, Jaime Carbonell, and Gregory Shreve

## Refereed Conference Publications

*The Sockeye 2 Neural Machine Translation Toolkit at AMTA 2020*, Tobias Domhan, Michael Denkowski, David Vilar, Felix Hieber, Xing Niu, and Kenneth Heafield. Proceedings of the Conference of the Association for Machine Translation in the Americas, 2020 (**AMTA 2020**)

*The Sockeye Neural Machine Translation Toolkit at AMTA 2018*, Felix Hieber, Tobias Domhan, Michael Denkowski, David Vilar, Artem Sokolov, Ann Clifton, and Matt Post. Proceedings of the Conference of the Association for Machine Translation in the Americas, 2018 (**AMTA 2018**)

*Cognitive Demand and Cognitive Effort in Post-Editing*, Isabel Lacruz, Michael Denkowski, and Alon Lavie. Proceedings of the Conference of the Association for Machine Translation in the Americas, 2014 (**AMTA 2014**)

*Learning from Post-Editing: Online Model Adaptation for Statistical Machine Translation*, Michael Denkowski, Chris Dyer, and Alon Lavie. Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, 2014 (**EACL 2014**)

*Analyzing and Predicting MT Utility and Post-Editing Productivity in Enterprise-scale Translation Projects*, Olga Beregovaya, Alon Lavie, David Clarke, and Michael Denkowski. Proceedings of the 12th Machine Translation Summit, 2013 (**MT Summit 2013**)

*Challenges in Predicting Machine Translation Utility for Human Post-Editors*, Michael Denkowski and Alon Lavie. Proceedings of the Conference of the Association for Machine Translation in the Americas, 2012 (**AMTA 2012**)

*Choosing the Right Evaluation for Machine Translation: an Examination of Annotator and Automatic Metric Performance on Human Judgment Tasks*, Michael Denkowski and Alon Lavie. Proceedings of the Conference of the Association for Machine Translation in the Americas, 2010 (**AMTA 2010**)

*Extending the METEOR Machine Translation Evaluation Metric to the Phrase Level*, Michael Denkowski and Alon Lavie. Proceedings of Human Language Technologies: the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, 2010 (**NAACL-HLT 2010**)

## Refereed Workshop Publications

*Bi-Directional Neural Machine Translation with Synthetic Parallel Data*, Xing Niu, Michael Denkowski, and Marine Carpuat. Proceedings of the ACL 2018 Workshop on Neural Machine Translation and Generation (**WNMT 2018**)

*Stronger Baselines for Trustable Results in Neural Machine Translation*, Michael Denkowski and Graham Neubig. Proceedings of the ACL 2017 Workshop on Neural Machine Translation (**WNMT 2017**)

*Real Time Adaptive Machine Translation for Post-Editing with cdec and TransCenter*, Michael Denkowski, Alon Lavie, Isabel Lacruz, and Chris Dyer. Proceedings of the EACL 2014 Workshop on Humans and Computer-assisted Translation, 2014 (**HaCaT 2014**)

*Meteor Universal: Language Specific Translation Evaluation for Any Target Language*, Michael Denkowski and Alon Lavie. Proceedings of the ACL 2014 Ninth Workshop on Statistical Machine Translation, 2014, (**WMT 2014**)

*The CMU Machine Translation Systems at WMT 2013: Syntax, Synthetic Translation Options, and Pseudo-References*, Waleed Ammar, Victor Chahuneau, Michael Denkowski, Greg Hanneman, Wang Ling, Austin Matthews, Kenton Murray, Nicola Segall, Alon Lavie and Chris Dyer. Proceedings of the ACL 2013 Eighth Workshop on Statistical Machine Translation, 2013 (**WMT 2013**)

*The CMU-Avenue French-English Translation System*, Michael Denkowski, Greg Hanneman, and Alon Lavie. Proceedings of the NAACL 2012 Seventh Workshop on Statistical Machine Translation, 2012 (**WMT 2012**)

*Meteor 1.3 Automatic Metric for Reliable Optimization and Evaluation of Machine Translation Systems*, Michael Denkowski and Alon Lavie. Proceedings of the EMNLP 2011 Sixth Workshop on Statistical Machine Translation, 2011, (**WMT 2011**)

*METEOR-NEXT and the METEOR Paraphrase Tables: Improved Evaluation Support For Five Target Languages*, Michael Denkowski and Alon Lavie. Proceedings of the ACL 2010 Joint Workshop on Statistical Machine Translation and Metrics MATR, 2010 (**WMT 2010**)

*Turker-Assisted Paraphrasing for English-Arabic Machine Translation*, Michael Denkowski, Hassan Al-Haj, and Alon Lavie. Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data With Amazon's Mechanical Turk, 2010

*Exploring Normalization Techniques for Human Judgments of Machine Translation Adequacy Collected Using Amazon Mechanical Turk*, Michael Denkowski and Alon Lavie. Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data With Amazon's Mechanical Turk, 2010

## Journal Articles

*The METEOR Metric for Automatic Evaluation of Machine Translation*, Alon Lavie and Michael Denkowski. Machine Translation, 2009

## Book Chapters

*Searching for Better Automatic MT Metrics*, Lavie, Agarwal, Denkowski, Snover, Madnani, Dorr, Schwartz, Habash, Kahn, Ostendorf, Roark, Kulick, Marcus, Pado, Galley, and Manning. Handbook of Natural Language Processing and Machine Translation, 2011

## Unrefereed Reports

*The Amazon Nova family of models: Technical report and model card*, Amazon Artificial General Intelligence. Amazon Technical Reports, 2024

*Sockeye 3: Fast Neural Machine Translation with PyTorch*, Felix Hieber, Michael Denkowski, Tobias Domhan, Barbara Darques Barros, Celina Dong Ye, Xing Niu, Cuong Hoang, Ke Tran, Benjamin Hsu, Maria Nadejde, Surafel Lakew, Prashant Mathur, Anna Currey, Marcello Federico. arXiv preprint arXiv:2207.05851, 2022

*Sockeye: A Toolkit for Neural Machine Translation*, Felix Hieber, Tobias Domhan, Michael Denkowski, David Vilar, Artem Sokolov, Ann Clifton and Matt Post. arXiv preprint arXiv:1712.05690, 2017

*TransCenter: Web-Based Translation Research Suite*, Michael Denkowski and Alon Lavie. Proceedings of the AMTA 2012 Workshop on Post-Editing Technology and Practice Demo Session, 2012

*METEOR-Tuned Phrase-Based SMT: CMU French-English and Haitian-English Systems for WMT 2011*, Michael Denkowski and Alon Lavie. CMU Technical Report CMU-LTI-11-011, Language Technologies Institute, Carnegie Mellon University, 2011

*A Survey of Techniques for Unsupervised Word Sense Induction*, Michael Denkowski. Literature Review for Language and Statistics II Course at Carnegie Mellon University, 2009

## Open Source Software

### Primary Developer

#### *Meteor*

Automatic machine translation evaluation system with state-of-the-art performance in several tasks  
Project page: <https://github.com/cmu-mtlab/meteor>

#### *TransCenter*

Web-based translation editing environment with support for personalized machine translation  
Project page: <https://github.com/cmu-mtlab/transcenter-live>

#### *qe-clean*

Automatic bilingual data cleaning inspired by machine translation quality estimation  
Project page: <https://github.com/cmu-mtlab/qe-clean>

### Contributor

#### *Sockeye*

Amazon's sequence-to-sequence modeling toolkit built on PyTorch  
Project page: <https://github.com/aws-labs/sockeye>

#### *Moses*

Statistical phrase-based machine translation system  
Project page: <http://www.statmt.org/ Moses>

#### *cdec*

Decoder, aligner, and learning framework for machine translation and structured prediction  
Project page: <http://www.cdec-decoder.org>

## Involvement

### Board Member

Texas Christian University Computer Science Industry Advisory Board: Member 2021-Present  
Association for Machine Translation in the Americas (AMTA): Research Co-Director 2019-2020

### Program Committee Member / Reviewer

Conference on Machine Translation (WMT): 2011-2024  
Association for Computational Linguistics (ACL): 2014, 2016-2021, 2023  
North American Chapter of the Association for Computational Linguistics (NAACL): 2013, 2021  
European Chapter of the Association for Computational Linguistics (EACL): 2014, 2017

Asia-Pacific Chapter of the Association for Computational Linguistics (ACL): 2022  
Empirical Methods in Natural Language Processing (EMNLP): 2012, 2014-2018, 2021-2023  
Conference on Computational Linguistics (COLING): 2016, 2018  
Conference of the Association for Machine Translation in the Americas (AMTA): 2014  
Conference of the European Association for Machine Translation (EAMT): 2020  
Workshop on Post-editing Technology and Practice (WPTP): 2015  
Journal of Artificial Intelligence Research (JAIR): 2022  
Conference on Language Modeling (COLM): 2024

### **Invited Talks**

*Adaptive Machine Translation Systems for Post-Editing Tasks*  
at the First Machine Translation Marathon in the Americas, May 2015  
*Learning from Post-Editing: Real Time Model Adaptation for Machine Translation*  
at the AMTA 2014 Workshop on Interactive and Adaptive Machine Translation, October 2014

### **Interns Mentored**

#### **Amazon.com, Inc.**

Xing Niu (University of Maryland, College Park): leveraging monolingual data for neural machine translation, Summer 2017  
Shuoyang Ding (Johns Hopkins University): automatic hyperparameter optimization for statistical machine translation, Summer 2016  
Kartik Goyal (Carnegie Mellon University): machine translation quality estimation, Summer 2015

Curriculum vitae updated February 2025