

RAMI AL-RFOU

rami@alrfou.com +1 631 371 3165
http://alrfou.com 701 N Rengstorff, Apt 19
Mountain View, CA 94043

Experience

- SEP 20 - *present* **Staff Research Scientist**
APR 17 - SEP 20 **Senior Research Scientist**
JUN 15 - APR 17 **Software Engineer**, *Google Research*, Mountain View, CA.
Technical lead for SmartReply. Worked on deep retrieval systems, conversational modeling, and query understanding.
- MAY - AUG 13 **Research Intern**, *Microsoft*, New York City, NY.
Worked with Léon Bottou on active learning and prior knowledge integration.
- MAY - AUG 12 **Research Intern**, *Google Research*, Mountain View, CA.
Designed a multilingual coreference resolution system with Jay Ponte.
- JUN 11 - AUG 11 **Software Engineering Intern**, *Google*, Mountain View, CA.
Worked on visualizing Google network infrastructure.
- FEB 09 - JULY 10 **Lab Instructor**, *Middle East Technical University*, TRNC.
APR 08 - DEC 08 **Software Developer**, *Jovision*, Amman, Jordan.

Selected Publications

Language Modeling

- Multilingual **mT5: A Massively Multilingual Pre-trained Text-to-Text Transformer.**
Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, Colin Raffel. *under review*.
- LAReQA: Language-agnostic answer retrieval from a multilingual pool.**
Uma Roy, Noah Constant, Rami Al-Rfou, Aditya Barua, Aaron Phillips, Yinfei Yang. EMNLP '20.
- Wiki-40B: Multilingual Language Model Dataset.**
Mandy Guo, Zihang Dai, Denny Vrandečić, Rami Al-Rfou. LREC '20.
- Polyglot-NER: Massive Multilingual Named Entity Recognition.**
Rami Al-Rfou, Vivek Kulkarni, Bryan Perozzi, and Steven Skiena. SDM '15, Vancouver, Canada.
- Token-Free **Character-Level Language Modeling with Deeper Self-Attention.**
Rami Al-Rfou, Dokook Choe, Noah Constant, Mandy Guo, Llion Jones. 33rd AACL, Honolulu, Hawaii.
- Bridging the Gap for Tokenizer-Free Language Models.**
Dokook Choe, Rami Al-Rfou, Mandy Guo, Heeyoung Lee, Noah Constant. BayLearn '19.
- Knowledge Base **Knowledge Graph Based Synthetic Corpus Generation for Knowledge-Enhanced Language Model Pre-training.**
Oshin Agarwal, Heming Ge, Siamak Shakeri, Rami Al-Rfou. *under-review*.
- Machine Translation Aided Bilingual Data-to-Text Generation and Semantic Parsing.**
Oshin Agarwal, Mihir Kale, Heming Ge, Siamak Shakeri, Rami Al-Rfou. 3rd Workshop on Natural Language Generation from the Semantic Web.
- Conversational **Efficient Natural Language Response Suggestion for Smart Reply.**
Matthew Henderson, Rami Al-Rfou, Brian Strope, Yun-hsuan Sung, Laszlo Lukacs, Ruiqi Guo
Sanjiv Kumar, Balint Miklos, Ray Kurzweil
- Conversational Contextual Cues: The Case of Personalization and History for Response Ranking**
Rami Al-Rfou, Marc Pickett, Javier Snaider, Yun-hsuan Sung, Brian Strope, Ray Kurzweil

Graph Neural Networks

- AUG 19 **DDGK: Learning Graph Representations for Deep Divergence Graph Kernels**
Rami Al-Rfou, Dustin Zelle, Bryan Perozzi. *WWW 2019, San Francisco, USA.*
- DEC 18 **Watch Your Step: Learning Node Embeddings via Graph Attention**
Sami Abu-El-Haija, Bryan Perozzi, Rami Al-Rfou, and Alexander A Alemi. *NeurIPS 2018, Montreal, Canada.*
- NOV 17 **Learning Edge Representations via Low-Rank Asymmetric Projections**
Sami Abu-El-Haija, Bryan Perozzi, and Rami Al-Rfou. *CIKM 2017, NYC, USA.*
- AUG 14 **DeepWalk: Online Learning of Social Representations.**
Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. *20th ACM SIGKDD 2014, NYC, USA.*

Patents

- US-2020334495-A1 Systems and Methods for Determining Graph Similarity.
- US-2019197101-A1 Selective text prediction for electronic messaging.
- US-2018240013-A1 Cooperatively training and/or using separate input and subsequent content neural networks for information retrieval.
- US-2018240014-A1 Cooperatively training and/or using separate input and response neural network models for determining response(s) for electronic communications.
- US-9514098-B1 Iteratively learning coreference embeddings of noun phrases using feature representations that include distributed word representations of the noun phrases.

Education

- SEP 10 - MAY 15 **PhD in Computer Science**, *Stony Brook University*, Stony Brook, NY.
Thesis: *Polyglot: Massive multilingual Natural Language Processing Pipeline.*
Adviser: *Steven Skiena.*
- SEP 04 - FEB 09 **BSc. in Computer Engineering**, *University of Jordan*, Amman, Jordan.
GPA: 3.79/4.0, ranked 4th of a 128-student class.

Voluntary Work

- 2018 - *present* **Supporter** of Manara's efforts to increase Arab engineers representation in the tech industry. (manara.tech)
- 2015 - *present* **Reviewer** for ACL, EMNLP, NAACL, LREC, TPAMI, DAMI, COLING, WebConf.
- OCT 07 - SEP 08 **Chairman**, *IEEE student branch*, University of Jordan, Amman, Jordan.
- JAN - JUNE 08 **Linux System Administrator**, *Theoretical Physics Lab*, University of Jordan.
- OCT 06 - JAN 08 **Team Leader**, *"We Are All Jordan" Youth Commission*, Jordan.
• Represented the Jordanian youth as a part of the official delegation during the state visit of His Majesty King AbduAllah II to China, from Oct.27th - Nov.1st 2007.