

# Ahmed El-Kishky

MEMBER OF TECHNICAL STAFF · OPEN AI

San Francisco, CA 94110

✉ ahmedelkishky09@gmail.com | 🏠 ahelk.github.io | 📧 ahelk | 🌐 ahelk

## Research Interests

- ◇ My interests are broadly in large-scale machine learning, natural language processing, graph machine learning, and recommender systems.

## Research and Industry Experience

### OpenAI

MEMBER OF TECHNICAL STAFF

San Francisco, CA

May 2023 - Present

- Teaching large language models to solve hard problems.

### Twitter

Seattle, WA

STAFF ML RESEARCHER - TECH LEAD

April 2021 - April 2023

- Proposed and developed the first in-house Twitter graph embeddings (TwHIN embeddings). These have demonstrated production metric impact across many recommendation verticals including Ads ranking (+7% revenue, +11% ads value pre-acquisition, +9.6% post-acquisition), Search ranking (+3% search engagement), Home Timeline ranking (+20m user active minutes), and notifications ranking (+1M daily monetizable users).
- Proposed and led the development of TwHIN-BERT, the first large language model at Twitter which we train with text self-supervision and user social engagements.

### Meta AI

Seattle, WA

SENIOR RESEARCH SCIENTIST

Sept 2018 - March 2021

- Research scientist working on large-scale data mining, and multi-lingual NLP. More specifically led efforts to facilitate massively multilingual transformer-based machine translation models, published papers, and worked with engineers to apply these NLP techniques to move company metrics (engagement and translation quality).

### Facebook

Seattle, WA

PHD INTERN

2017

- Performed NLP research in supervised topic modeling. Applied this work to classify unstructured text into content topics for recommender systems. Also worked on intent prediction on unstructured content for ads targeting.
- Performed research into multimodal representation learning (joint text + image embedding).

### Facebook

Menlo Park, CA

PHD INTERN

2016

- Performing machine learning modeling for search ranking. Developed an association rule mining that identifies rules between queries and documents to improve candidate retrieval for search.

### Google Research

Mountain View, CA

RESEARCH INTERN

2016

- Performed research into scalable and distributed techniques for generating higher-order cross features for machine learning.

### Palantir Technologies

Palo Alto, CA

MACHINE LEARNING SOFTWARE ENGINEER

Summer 2015

- Developed time-series anomaly detection techniques with an application to high-volume sensor data.

### Microsoft Research

Redmond WA

RESEARCH INTERN

Spring 2015

- Developed an information retrieval and data mining approach to leverage query logs to provide a low-code interactive interface for automatically generating SQL queries.

### Rocket Fuel Inc

Redwood City, CA

SOFTWARE ENGINEERING INTERN

Summer 2013

- Performed data science and software engineering work within the mobile advertising team.

## University of Massachusetts Amherst

Amherst, MA

NSF REU RESEARCHER

Summer 2012

- Performed research and development into Bayes net structure learning, causality discovery, and unsupervised statistical clustering.

## University of Notre Dame

Notre Dame, IN

NSF REU RESEARCHER

Summer 2011

- Performed research and development into leveraging high-performance computing with sensor networks for modeling fluid dynamics (e.g., Navier Stokes & Advection Diffusion)

## Education

---

### University of Illinois at Urbana-Champaign

PH.D. IN COMPUTER SCIENCE

- Advisor: Jiawei Han
- GPA: 4.0

### University of Tulsa

B.S. IN COMPUTER SCIENCE & B.S. IN MATHEMATICS

- GPA: 3.91

## Publications & Pre-prints

---

- ◇ J. Pougué-Biyong, A. Gupta, A. Haghighi, A.El-Kishky, “[Learning Stance Embeddings from Signed Social Graphs](#)”, *Proceedings of WSDM 2023* **Oral Presentation Best Paper Award**
- ◇ X. Zhang, Y. Malkov, O. Florez, S. Park, B. McWilliams, J. Han, A. El-Kishky, “[TwHIN-BERT: A Socially-Enriched Pre-trained Language Model for Multilingual Tweet Representations](#)”, *Proceedings of SIGKDD 2023* **Oral Presentation**
- ◇ A. El-Kishky, T. Markovich, K. Leung, F. Portman, A. Haghighi, Y. Xiao, “[kNN-Embed: Locally-Smoothed Embedding Mixtures for Diverse Candidate Retrieval](#)”, *Proceedings of PAKDD 2023* **Oral Presentation**
- ◇ A. El-Kishky, T. Markovich, S. Park, C. Verma, B. Kim, R. Eskander, Y. Malkov, F. Portman, S. Samaniego, Y. Xiao, A. Haghighi, “[TwHIN: Embedding the Twitter Heterogeneous Information Network for Personalized Recommendation](#)”, *Proceedings of SIGKDD 2022* **Oral Presentation**
- ◇ F. Mireshghallah, N. Vogler, J. He, O. Florez, A. El-Kishky, T. Berg-Kirkpatrick, “[Non-Parametric Temporal Adaptation for Social Media Topic Classification](#)”, *arXiv* **Preprint**
- ◇ F. Portman, S. Ragain, A. El-Kishky, “[MiCRO: Multi-interest Candidate Retrieval Online](#)”, *arXiv* **Preprint**
- ◇ J. Li, S. Mishra, A. El-Kishky, S. Mehta, V. Kulkarni, “[NTULM: Enriching Social Media Text Representations with Non-Textual Units](#)”, *WNUT* **Oral Presentation**
- ◇ S. Sun, A. El-Kishky, V. Chaudhary, J. Cross, F. Guzman, L. Specia, “[Classification-based Quality Estimation: Small and Efficient Models for Real-world Applications](#)”, *Proceedings of EMNLP 2021* **Poster**
- ◇ A. El-Kishky, J. Cross, A. Renduchintala, F. Guzman, P. Koehn, “[XLEnt: Mining Cross-lingual Entities with Lexical-Semantic-Phonetic Word Alignment](#)”, *Proceedings of EMNLP 2021* **Oral Presentation**
- ◇ J. Wang, C. Xu, F. Guzman, A. El-Kishky, B. Rubinstein, T. Cohn, “[As Easy as 1, 2, 3: Behavioural Testing of NMT Systems for Numerical Translation](#)”, *Findings of ACL 2021* **Journal**
- ◇ C. Xu, J. Wang, F. Guzman, A. El-Kishky, Y. Tang, B. Rubinstein, T. Cohn, “[Putting Words Into the System’s Mouth: A Targeted Attack on Neural Machine Translation Using Monolingual Data Poisoning](#)”, *Findings of ACL 2021* **Journal**
- ◇ W. Ko, A. El-Kishky, A. Renduchintala, V. Chaudhary, N. Goyal, F. Guzman, P. Fung, M. Diab, P. Koehn, “[Adapting High-resource NMT Models to Translate Low-resource Related Languages without Parallel](#)

[Data](#)”, *Proceedings of ACL 2021* **Oral Presentation**

- ◇ A. Fan, S. Bhosale, H. Schwenk, Z. Ma, A. El-Kishky, S. Goyal, M. Baines, O. Celebi, G. Wenzek, V. Chaudhary, N. Goyal, T. Birch, V. Liptchinsky, S. Edunov, E. Grave, M. Auli, A. Joulin, “[Beyond English-Centric Multilingual Machine Translation](#)”, *Journal of Machine Learning Research (JMLR) 2021* **Journal**
- ◇ Y. Tuan, A. El-Kishky, A. Renduchintala, V. Chaudhary, F. Guzman, L. Specia, “[Quality Estimation without Human-labeled Data](#)”, *Proceedings of EACL 2021* **Poster**
- ◇ P. Koehn, V. Chaudhary, A. El-Kishky, N. Goyal, P. Chen, F. Guzman, “[Findings of the WMT 2020 Shared Task on Parallel Corpus Filtering](#)”, *Fifth Conference on Machine Translation (WMT’20)*
- ◇ A. El-Kishky, V. Chaudhary, F. Guzman, P. Koehn, “[CCAligned: A Massive Collection of Cross-lingual Web-Document Pairs](#)”, *Proceedings of EMNLP 2020* **Poster**
- ◇ S. Sun, M. Fomicheva, F. Blain, V. Chaudhary, A. El-Kishky, A. Renduchintala, F. Guzman, L. Specia, “[An Exploratory Study on Multilingual Quality Estimation](#)”, *AACL-IJCNLP 2020* **Oral Presentation**
- ◇ A. El-Kishky and F. Guzman, “[Massively Multilingual Document Alignment with Cross-lingual Sentence-Mover’s Distance](#)”, *Proceedings of AACL-IJCNLP 2020* **Oral Presentation**
- ◇ D. Golub, R. Martín-Martín, A. El-Kishky, S. Savarese, “[Leveraging Pretrained Image Classifiers for Language-Based Segmentation](#)”, *Winter Conference on Applications of Computer Vision 2020* **Oral Presentation**
- ◇ P. Chen, J. Shen, M. Le, V. Chaudhary, A. El-Kishky, G. Wenzek, M. Ott, M. Ranzato., “[Facebook AI’s WAT19 Myanmar-English Translation Task Submission](#)”, *WAT 2019* **Workshop**
- ◇ A. El-Kishky, X. Fu, A. Addawood, N. Sobh, C. Voss, J. Han, “[Constrained Sequence-to-sequence Semitic Root Extraction for Enriching Word Embeddings](#)”, *In Proceedings of the Fourth Arabic Natural Language Processing Workshop, 2019.* **Workshop**
- ◇ H. Kim, A. El-Kishky, X. Ren, J. Han, “[Mining News Events from Comparable News Corpora: A Multi-Attribute Proximity Network Modeling Approach](#)”, *Proceedings of IEEE BigData 2019.* **Oral Presentation**
- ◇ A., El-Kishky, F. Xu, A. Zhang, J. Han, “[Parsimonious Morpheme Segmentation with an Application to Enriching Word Embeddings.](#)”, *Proceedings of IEEE BigData 2019.* **Oral Presentation**
- ◇ Q. Zhu, X. Ren, J. Shang, Y. Zhang, A. El-Kishky, J. Han, “[Integrating local context and global cohesiveness for open information extraction.](#)”, *Proceedings of WSDM 2019* **Oral Presentation**
- ◇ D. Xin, A. El-Kishky, D. Liao, B. Norick, J. Han, “[Active Learning on Heterogenous Information Networks: A Multi-armed Bandit Approach](#)”, *Proceedings of ICDM 2018.* **Oral Presentation**
- ◇ A. El-Kishky, F. Xu, A. Zhang, S. Macke, J. Han, “[Entropy-Based Subword Mining for Word Embeddings](#)”, *SCLeM Workshop at NAACL’2018* **Workshop**
- ◇ A. El-Kishky, Y. Song, C. Wang, C. Voss, J. Han, “[Scalable Topical Phrase Mining from Text Corpora](#)”, *Proceedings of VLDB 2015* **Oral Presentation**
- ◇ X. Ren, A. El-Kishky, C. Wang, F. Tao, C. Voss, H. Ji, J. Han, “[ClusType: Effective Entity Recognition and Typing by Relation Phrase-Based Clustering](#)”, *Proceedings of SIGKDD 2015* **Oral Presentation**
- ◇ C. Wang, Y. Song, A. El-Kishky, D. Roth, M. Zhang, J. Han, “[Incorporating World Knowledge to Document Clustering via Heterogeneous Information Networks](#)”, *Proceedings of SIGKDD 2015* **Oral Presentation**
- ◇ F. Tao, J. Han, H. Ji, G. Brova, C. Wang, A. El-Kishky, J. Liu, X. Ren, Y. Sun, “[NewsNetExplorer: Automatic Construction and Exploration of News Information Networks](#)”, *Proceedings of SIGMOD 2014* **Demo**
- ◇ D. Arbour, J. Atwood, A. El-Kishky, D. Jensen, “[Agglomerative Clustering of Bagged Data Using Joint Distributions](#)”, *In Papers from ICML: SLG 2013* **Workshop**
- ◇ A. El-Kishky, “[Assessing Entropy and Fractal Dimensions as Discriminants of Seizures in EEG Time Series](#)”, *Information Science, Signal Processing and their Applications (ISSPA 2012)* **Oral Presentation**

## Conference Tutorial Presentations

---

- ◇ A. El-Kishky, M. Bronstein, Y. Xiao, & A. Haghighi, “[Graph-based Representation Learning for Web-scale Recommender Systems](#)”, *ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (SIGKDD 2022)*
- ◇ A. El-Kishky, M. Bronstein, Y. Xiao, & A. Haghighi, “[Graph Embeddings for Web-scale Recommender Systems](#)”, *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD 2022)*
- ◇ A. El-Kishky, P. Koehn, & H. Schwenk, “[Searching the Web for Cross-lingual Parallel Data](#)”, *ACM SIGIR Conf. on Information Retrieval (SIGIR 2020)*
- ◇ X. Ren, A. El-Kishky, H. Ji, & J. Han, “[Automatic Entity Recognition and Typing in Massive Text Data](#)”, *ACM SIGMOD Conf. on Management of Data (SIGMOD 2016)*
- ◇ X. Ren, A. El-Kishky, C. Wang, & J. Han, “[Automatic Entity Recognition and Typing in Massive Text Corpora](#)”, *International World Wide Web Conference (WWW 2016)*
- ◇ X. Ren, A. El-Kishky, C. Wang, & J. Han, “[Automatic Entity Recognition and Typing from Massive Text Corpora: A Phrase and Network Mining Approach](#)”, *ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (SIGKDD 2015)*
- ◇ J. Han, C. Wang, A. El-Kishky, “[Bringing Structure to Text: Mining Phrases, Entities, Topics, and Hierarchies](#)”, *ACM SIGKDD Conf. on Knowledge Discovery and Data Mining (SIGKDD 2014)*

## Honors & Awards

---

- 2023 **Best Paper**, 16th ACM International Conference on Web Search and Data Mining (WSDM)<sup>†</sup>
- 2016 **Awardee**, State Farm Doctoral Fellowship\*
- 2015 **Awardee**, NDSEG Graduate Research Fellowship\*
- 2013 **Awardee**, NSF Graduate Research Fellowship\*
- 2013 **1st Place**, Team leader in DWR Governor’s Cup Entrepreneurship Competition (\$22,000 prize)
- 2012 **3rd Place**, Team leader in DWR Governor’s Cup Entrepreneurship Competition (\$6,000 prize)
- **1st Place**, MAA Oklahoma/Arkansas Mathematics Competition (2011,2012,2013)

## Skills

---

<b>Programming Languages</b>	Python, Java, C#, C/C++, SQL, $\LaTeX$
<b>Machine Learning Tools</b>	PyTorch, NumPy, SciPy
<b>Distributed Tools</b>	Map Reduce, Hive, Spark, Flume
<b>Spoken Languages</b>	English (Native), Arabic (Native)

## Students Supervised

---

<b>Yi-lin Tuan</b>	PhD University of California Santa Barbara
<b>Shuo Sun</b>	PhD Johns Hopkins University
<b>Wei-Jen Ko</b>	PhD University of Texas at Austin
<b>Xingyu Fu</b>	PhD University of Pennsylvania
<b>John Pougué-Biyong</b>	PhD University of Oxford
<b>Tianyang Zhao</b>	PhD University of California Los Angeles
<b>Xinyang Zhang</b>	PhD University of Illinois at Urbana-Champaign

---

\*Nationally competitive

†Internationally competitive