Ehsan Kamalloo

Homepage: https://ehsk.github.io/   Email: ekamalloo@uwaterloo.ca

Professional Experience

Jan 2023 - *Now*	**Post-doctoral Fellow**, hosted by Jimmy Lin, David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, Canada.
Mar 2021 - Dec 2022	**Associate Researcher (part-time)**, Huawei Noah's Ark, Montreal, Canada.
Sep 2020 - Mar 2021	**Research Intern**, Huawei Noah's Ark, Montreal, Canada.

Education

2016 - 2022	**Ph.D. Computing Science**, University of Alberta, Edmonton, Canada. Thesis: "Robust Knowledge Acquisition in Answering Information-seeking Questions At Scale," Supervised by Davood Rafiei GPA: 4.0/4.0
2008 - 2011	**M.Sc. Software Engineering**, Tarbiat Modares University, Tehran, Iran.
2004 - 2008	**B.Sc. Software Engineering**, University of Tehran, Tehran, Iran.

Selected Publications

Preprints

[1] **Ehsan Kamalloo**, Aref Jafari, Xinyu Zhang, Nandan Thakur, and Jimmy Lin (2023), "HAGRID: A Human-LLM Collaborative Dataset for Generative Information-seeking with Attribution."

[2] **Ehsan Kamalloo**, Nandan Thakur, Carlos Lassance, Xueguang Ma, Jheng-Hong Yang, and Jimmy Lin (2023), "Resources for Brewing BEIR: Reproducible Reference Models and an Official Leaderboard."

Conference Proceedings and Journals

[1] **Ehsan Kamalloo**, Nouha Dziri, Charles Clarke, and Davood Rafiei (2023), "Evaluating Open-Domain Question Answering in the Era of Large Language Models," in *ACL (oral)*.

[2] **Ehsan Kamalloo**, Xinyu Zhang, Odunayo Ogundepo, Nandan Thakur, David Alfonso-Hermelo, Mehdi Rezagholizadeh, and Jimmy Lin (2023), "Evaluating Embedding APIs for Information Retrieval," in *ACL Industry Track*.

[3] **Ehsan Kamalloo**, Charles Clarke, and Davood Rafiei (2023), "Limitations of Open-Domain Question Answering Benchmarks for Document-level Reasoning," in *SIGIR*.

[4] Xinyu Zhang, Nandan Thakur, Odunayo Ogundepo, **Ehsan Kamalloo**, David Alfonso-Hermelo, Xiaoguang Li, Qun Liu, Mehdi Rezagholizadeh, and Jimmy Lin (2023), "Making a MIRACL: Multilingual Information Retrieval Across a Continuum of Languages," in *TACL*.

[5] Nouha Dziri, **Ehsan Kamalloo**, Sivan Milton, Osmar Zaiane, Mo Yu, Edoardo M. Ponti, and Siva Reddy (2022), "FAITHDIAL: A Faithful Benchmark for Information-Seeking Dialogue," in *TACL*.

[6] Mehdi Akbarian Rastaghi, **Ehsan Kamalloo**, and Davood Rafiei (2022), "Probing the Robustness of Pre-trained Language Models for Entity Matching," in *CIKM*.

[7] **Ehsan Kamalloo**, Mehdi Rezagholizadeh, and Ali Ghodsi (2022), "When Chosen Wisely, More Data Is What You Need: A Universal Sample-Efficient Strategy For Data Augmentation," in *ACL Findings*.

[8] Mostafa Yadegari, **Ehsan Kamalloo**, and Davood Rafiei (2022), "Detecting Frozen Phrases in Open-Domain Question Answering," in *SIGIR*.

[9] **Ehsan Kamalloo**, Mehdi Rezagholizadeh, Peyman Passban, and Ali Ghodsi (2021), "Not Far Away, Not So Close: Sample Efficient Nearest Neighbour Data Augmentation via MiniMax," in *ACL Findings*.

[10] Nouha Dziri, **Ehsan Kamalloo**, Kory W. Mathewson, and Osmar R. Zaiane (2019), "Evaluating Coherence in Dialogue Systems using Entailment," in *NAACL*.

[11] Nouha Dziri, **Ehsan Kamalloo**, Kory W. Mathewson, and Osmar R. Zaiane (2019), "Augmenting Neural Response Generation with Context-Aware Topical Attention," in *ACL Workshop on NLP for Conversational AI (NLP4ConvAI)*.

[12] **Ehsan Kamalloo** and Davood Rafiei (2018), "A coherent unsupervised model for toponym resolution," in *World Wide Web (WWW) Conference*, pp. 1287-1296.

[13] Abolfazl AleAhmad, **Ehsan Kamalloo**, Arash Zareh, Masoud Rahgozar, and Farhad Oroumchian (2008), "Cross language experiments at Persian@ CLEF 2008," in *Cross Language Evaluation Forum (CLEF)*, pp. 105-112.

ACADEMIC SERVICE

Reviewing:

- **Conferences**: ACL Rolling Review, EMNLP 2023, CIKM 2023, ACL 2023, EACL 2023, EMNLP 2022, EMNLP 2021, ACL 2021, NAACL 2021, EACL 2021, EMNLP 2020, ACL 2020.

- **Workshops**: DialDoc Workshop at ACL 2023, Efficient Natural Language and Speech Processing at NeurIPS 2022 & 2021, SustaiNLP at EMNLP 2021, Student Research Workshop at NAACL 2021 and EACL 2021, SustaiNLP at EMNLP 2020, Student Research Workshop at ACL 2019, 2020, 2023.

Organizing:

- Multilingual Information Retrieval Across a Continuum of Languages (MIRACL), WSDM Cup 2023.

TEACHING EXPERIENCE

Student Advisor: *Master's thesis*: Mostafa Yadegari and Mehdi Akbarian (University of Alberta). *Undergrad Interns*: Boris Fleysher and Peter Elliott (University of Alberta).

Mar 2022	**Guest Lecturer** for *Introduction to Computational Natural Language Processing*, McMaster University, Canada.
Nov 2021	**Guest Lecturer** for *Foundation Models in Natural Language Processing*, University of Tehran, Iran.

2018 - 2019	**Teaching Assistant** for *Database Management Systems*.
2016 - 2018	**Teaching Assistant** for *Introduction to File and Database Management*.

INDUSTRIAL EXPERIENCE

2016 2013	Director of Research and Development at ODDRUN startup, Tehran, Iran Led technical strategy, and managed R&D team, Developed a recommender system in a location-based social network with ∼200K users,

2013 2010	Software Architect and Senior Software Engineer at FANAP, Tehran, Iran Led a 4-person technical team in the insurance solution division,

2010 2008	Full-Stack Software Engineer at FANAP, Tehran, Iran Developed property insurance, life insurance and accounting software, components of an enterprise integrated insurance solution, deployed at nation-wide scale.