

# EHSAN KAMALLOO

Homepage: <https://cs.ualberta.ca/~kamalloo> GitHub: <https://github.com/ehsk>  
Email: [kamalloo@ualberta.ca](mailto:kamalloo@ualberta.ca) Phone: +1 (780) 707-8820

## 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.  
Thesis: “Credit Management Using Artificial Immune Systems”
- 2004 - 2008 **B.Sc. Software Engineering**, University of Tehran, Tehran, Iran.  
Thesis: “A Morphology-based Stemmer for the Persian language”

## RESEARCH EXPERIENCE

---

- |            |  |  |
|------------|--|--|
| SEP 2022   |  | <b>Research Assistant</b> , University of Alberta, Edmonton, Canada.   |
| SEP 2016   |  | Explore information retrieval models suitable for NLP tasks.   |
| <i>Now</i> |  |  |
| MAR 2021   |  | <b>Associate Researcher (part-time)</b> , Huawei Noah’s Ark, Montreal, Canada.<br>Work on sample efficiency in NLP and multilingual retrieval. |
| MAR 2021   |  | <b>Research Intern</b> , Huawei Noah’s Ark, Montreal, Canada.  |
| SEP 2020   |  | Worked on sample efficiency and data augmentation in NLP.  |

## PUBLICATIONS

---

### Preprints

- [1] Xinyu Zhang, Nandan Thakur, Odunayo Ogundepo, **Ehsan Kamalloo**, David Alfonso-Hermelo, Xiaoguang Li, Qun Liu, Mehdi Rezagholizadeh, and Jimmy Lin (2022), “Making a MIRACL: Multilingual Information Retrieval Across a Continuum of Languages.”
- [2] **Ehsan Kamalloo**, Charles Clarke, and Davood Rafiei (2022), “Document-level Reasoning: A Hidden Challenge in Open-domain Question Answering Benchmarks.”

### Conference Proceedings and Journals

- [1] 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*.
- [2] Mehdi Akbarian Rastaghi, **Ehsan Kamalloo**, and Davood Rafiei (2022), “Probing the Robustness of Pre-trained Language Models for Entity Matching,” in *CIKM*.

- [3] **Ehsan Kamaloo**, 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*.
- [4] Mostafa Yadegari, **Ehsan Kamaloo**, and Davood Rafiei (2022), “Detecting Frozen Phrases in Open-Domain Question Answering,” in *SIGIR*.
- [5] **Ehsan Kamaloo**, 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*.
- [6] Nouha Dziri, **Ehsan Kamaloo**, Kory W. Mathewson, and Osmar R. Zaiane (2019), “Evaluating Coherence in Dialogue Systems using Entailment,” in *NAACL*.
- [7] Nouha Dziri, **Ehsan Kamaloo**, 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)*.
- [8] **Ehsan Kamaloo** and Davood Rafiei (2018), “A coherent unsupervised model for toponym resolution,” in *World Wide Web (WWW) Conference*, pp. 1287-1296.
- [9] **Ehsan Kamaloo** and Mohammad Saniee Abadeh (2014), “Credit risk prediction using fuzzy immune learning,” in *Advances in Fuzzy Systems*, 2014: 7.
- [10] **Ehsan Kamaloo** and Mohammad Saniee Abadeh (2010), “An artificial immune system for extracting fuzzy rules in credit scoring,” in *IEEE Congress on Evolutionary Computation (CEC)*, pp. 1-8.
- [11] Abolfazl AleAhmad, **Ehsan Kamaloo**, 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, EACL 2023, EMNLP 2022, EMNLP 2021, ACL 2021, NAACL 2021, EACL 2021, EMNLP 2020, ACL 2020.
- **Workshops:** 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.

### Organizing:

- Multilingual Information Retrieval Across a Continuum of Languages (MIRACL), WSDM 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	<b>Guest Lecturer</b> for <i>Introduction to Computational Natural Language Processing</i> , McMaster University, Canada.
Nov 2021	<b>Guest Lecturer</b> for <i>Foundation Models in Natural Language Processing</i> , University of Tehran, Iran.
2018 - 2019	<b>Teaching Assistant</b> for <i>Database Management Systems</i> .
2016 - 2018	<b>Teaching Assistant</b> for <i>Introduction to File and Database Management</i> .

## PROFESSIONAL EXPERIENCE

---

2016	Director of Research and Development at ODDRUN startup, Tehran, Iran
2013	Led technical strategy, and managed R&D team, Developed a recommender system in a location-based social network with ~200K users,
2013	Software Architect and Senior Software Engineer at FANAP, Tehran, Iran
2010	Led a 4-person technical team in the insurance solution division,
2010	Full-Stack Software Engineer at FANAP, Tehran, Iran
2008	Developed property insurance, life insurance and accounting software, components of an enterprise integrated insurance solution, deployed at nation-wide scale.

## TECHNICAL SKILLS

---

Programming Language:	Python, Java, Scala, C/C++, Lua
ML Framework:	PyTorch, TensorFlow
Database:	Oracle, MySQL, PostgreSQL, SQLite
Framework:	Spark, Hadoop, Elasticsearch, Lucene, Redis, Kafka