

# Behzad Golshan

---

Research Scientist  
Megagon Labs (formerly Recruit Institute of Technology)  
444 Castro St, Suite 900  
Mountain View, CA 94041

**Email** behzad@megagon.ai  
behzad@bu.edu  
**Homepage** behzad.io

## Short Bio & Research Interests

---

I am currently a Research Scientist at *Megagon Labs* working on Natural Language Processing, Machine Learning, and Data Mining. I received my PhD from *Boston University* where I was a member of the *Data Management Group* in the Computer Science Department. I was fortunate to have Professor Evimaria Terzi as my advisor.

## Education

---

**2010-2016** PhD in Computer science - Boston University  
Advisor: Prof. Evimaria Terzi

**2005-2010** B.Sc. in Computer Engineering - University of Tehran  
Advisor: Prof. Masoud Asadpour

## Publications

---

1. S. Galhotra, B. Golshan, W. Tan, *Adaptive Rule Discovery for Labeling Text Data*, SIGMOD 2021.
2. J. Bjerva, N. Bhutani, B. Golshan, W. Tan, I. Augenstein, *SubjQA: A Dataset for Subjectivity and Review Comprehension*, EMNLP 2020
3. N. Bhutani, A. Traylor, C. Chen, X. Wang, B. Golshan, W. Tan, *Sampo: Unsupervised Knowledge Base Construction for Opinions and Implications*, AKBC 2020.
4. W. Hirota, Y. Suhara, B. Golshan, W. Tan, *Emu: Enhancing Multilingual Sentence Embeddings with Semantic Specialization*, AAI 2020.
5. A. Traylor, C. Chen, B. Golshan, X. Wang, Y. Li, Y. Suhara, J. Li, C. Demiralp, W. Tan, *Enhancing Review Comprehension with Domain-Specific Commonsense*, arXiv preprint arXiv:2004.03020.
6. B. Golshan, G. Mihaila, C. Chen, J. Engel, Al. Halevy, Y. Suhara, W. Tan, M. Matuschek, *Building a Hotel Concierge Bot: an industrial case study*, CAST 2019.
7. D. Ma, C. Chen, B. Golshan, W. Tan, *Essentia: Mining Domain-specific Paraphrases with Word-Alignment Graphs*, TextGraphs 2019.
8. B. Golshan, E. Terzi, P. Tsaparas, *Recommendations for optimizing the collective user experience*, SDM 2019.
9. S. Bahargam, B. Golshan, T. Lappas, E. Terzi, *A Team-Formation Algorithm for Faultline Minimization*, Expert Systems with Applications 2019.
10. X. Wang, J. Komiya, Y. Suhara, A. Feng, B. Golshan, A. Halevy, W. Tan, *Koko: A system for scalable semantic querying of text*, PVLDB 2018 (Demo Track).
11. X. Wang, A. Feng, B. Golshan, A. Halevy, G. Mihaila, H. Oiwa, W. Tan, *Scalable semantic querying of text*, PVLDB 2018.
12. A. Asai, S. Evensen, B. Golshan, A. Halevy, V. Li, A. Lopatenko, D. Stepanov, Y. Suhara, W. Tan, Y. Xu, *HappyDB: A Corpus of 100,000 Crowdsourced Happy Moments*, LREC 2018, Miyazaki, Japan.

13. C. Chen, B. Golshan, A. Halevy, W. Tan, A. Doan, *BigGorilla: An Open-Source Ecosystem for Data Preparation and Integration*, IEEE Data Engineering Bulletin 2018.
14. B. Golshan, E. Terzi, *Minimizing tension in teams*, CIKM 2017, Singapore.
15. R. Agrawal, B. Golshan, E. Papalexakis, *Homogeneity in web search results: Diagnosis and mitigation*, ACM TIST 2017.
16. E. Galbrun, B. Golshan, A. Gionis, E. Terzi, *Finding low-tension communities*, SDM 2017, Houston, USA.
17. B. Golshan, A. Halevy, G. Mihaila, W.C. Tan, *Data integration: after the teenage years*, SIGMOD/PODS 2017 (Invited paper), Chicago, USA.
18. R. Agrawal, B. Golshan, E. Papalexakis, *Overlap in the web search results of google and bing*, Journal of Web Science 2016.
19. R. Agrawal, B. Golshan, E. Papalexakis, *Toward Data-Driven Design of Educational Courses: A Feasibility Study*, JEDM 2016.
20. R. Agrawal, B. Golshan, E. Papalexakis, *Overlap between google and bing web search results!: twitter to the rescue?*, ACM COSn 2015, Stanford University, California, USA.
21. R. Agrawal, B. Golshan, E. Papalexakis, *A study of distinctiveness in web results of two search engines*, WWW 2015 Companion (Web Science Track), Florence, Italy.
22. R. Agrawal, B. Golshan, E. Papalexakis, *Whither social networks for web search?*, ACM SIGKDD 2015 (Industry Track), Sydney, Australia.
23. R. Agrawal, B. Golshan, E. Terzi, *Grouping students in educational settings*, ACM SIGKDD 2014, New York, USA.
24. B. Golshan, T. Lappas, E. Terzi, *Profit-maximizing cluster hires*, ACM SIGKDD 2014, New York, USA.
25. R. Agrawal, B. Golshan, E. Terzi, *Forming beneficial teams of students in massive online classes*, ACM L@S 2014, Work in Progress, Atlanta, USA.
26. B. Golshan, E. Terzi, *Unveiling variables in systems of linear equations*, SDM 2014, Philadelphia, USA.
27. B. Golshan, J. W. Byers, E. Terzi, *What do row and column marginals reveal about your dataset?*, NeurIPS 2013, Lake Tahoe, USA.
28. G. Misra, B. Golshan, E. Terzi, *A framework for evaluating the smoothness of data-mining results*, ECML-PKDD 2012, Bristol, UK.
29. B. Golshan, T. Lappas, E. Terzi, *Sofia Search: a tool for automating related-work search*, ACM SIGMOD 2012 (Demo Track), Scottsdale, USA.

## Research Experience

---

**Fall 16-Present** Research Scientist at Megagon Labs (formerly Recruit Institute of Technology)

**Summer 2015** Research Intern at Aalto University  
Host: Prof. Aristides Gionis

**Summer 2014** Research Intern at Microsoft Research (Silicon Valley)  
Host: Dr. Rakesh Agrawal

**Summer 2013** Research Intern at oDesk (now Upwork)  
Host: Dr. Panagiotis Papadimitriou

**Summer 2012** Research Intern at University di Roma la Sapienza  
Host: Prof. Aris Anagnostopoulos

**Summer 2011** Research Intern at Adverplex (now Cogo Labs)  
Host: Prof. Giorgos Zervas

## Honors & Awards

---

- Teaching Fellow Excellence Award – College of Arts and Sciences at Boston University (2012)
- Excellent Student M.Sc. Admission Offer to Tehran University (2009)
- Top 0.1% of the Iran's Nationwide Matriculation Exam (2005)
- 49th Rank in National Graduate Entrance Exam (2009)
- Faculty of Engineering Award (Tehran University) (2007)
- NSF travel award for KDD (2014)
- SDM travel award (2014)
- SIGMOD/PODS travel award (2012)

## Teaching Experience

---

### Teaching Assistant (Boston University):

- Combinatorial Structures (Spring 2012 and Spring 2013)
- Introductions to Computer Science I (Fall 2010)

### Teaching Assistant (Tehran University):

- Introductions to Computers and Programming (Fall 2008 and Fall 2009)
- Data Structures (Fall 2009)
- Algorithm Analysis and Design (Fall 2008)
- Principles of Compiler Design (Spring 2008)

## Professional Activities

---

### Organizing Committee:

- WIT: Workshop On Deriving Insights From User-Generated Text @ KDD2021 (2021)

### PC member and Reviewer:

- TKDE, NeurIPS, NLP4PI Workshop (2021)
- AAAI (2020)
- KDD (2019)
- ICML, CIKM, NeurIPS, ICDE (2018)
- NeurIPS, WSDM (2017)
- ICML, WSDM (2016)

### External Reviewer:

- Computational Intelligence (2016)
- TKDD, WSDM (2015)
- AAAI, WSDM, WWW (2014)
- ICDM, WWW, PODS, WSDM (2013)

### Volunteer:

- Student volunteer at IC2S2 (Finland) (2015)