

# Yingtao Tian

---

Computer Science Department, State University of New York at Stony Brook  
*E-mail:* alan.yt.tian@gmail.com

## EDUCATION

- State University of New York at Stony Brook**, New York, United States  
Ph.D., Computer Science Fall 2014 - Ongoing
- Advisor: [Prof. Steven Skiena](#)
  - Research interest: Natural language processing, especially on representation learning and multilingual applications, by combining deep learning and linguistics models.
- Fudan University**, Shanghai, China  
B.Sc., Computer Science and Technology 2010 - 2014

## EXPERIENCE

- Facebook**, California, United States  
*Software Engineering Intern* May 2017 - Aug. 2017
- Host: [Dr. Haixun Wang](#)
  - **Research on natural language understanding.** Researched on Web Query Parsing by inferring query treebank, and Natural Language Interface to Database System leveraging database table schema with Seq2seq models.
- Google Brain**, California, United States  
*Software Engineering Intern* May 2016 - Nov. 2016
- Host: [Dr. Stephan Gouws](#)
  - **Research on representation learning**  
Researched on investigating new representation learning techniques for both units (words) and sequences (sentences). Worked on sequence model for sentence representation and hierarchical embeddings.
- Google**, New York, United States  
*Software Engineering Intern, Search Group* May 2015 - Aug. 2015
- Host: [Dr. Xiaomeng Ban](#)
  - **User-generated activity profiling and analysis**  
Implemented a profiling tool for analyzing activities generated by billions of users. Built a full-stack framework using a proper combination of distributed processing, aggregation and visualization for effective large-scale analysis of user-generated content. Leveraged module-based design for flexibilities to easily integrate new profiling tasks.
- Microsoft Research Asia**, Beijing, China  
*Research Intern, Data Management, Analytics and Services Group* Feb. 2013 - July 2013
- Advisor: [Dr. Haixun Wang](#)
  - **Learning Semantic Vector Representation from Large Knowledge Base**  
Proposed and made experiments on a framework generating semantic embeddings (vector representation) from Probbase, a large knowledge base and probabilistic taxonomy.
- Research Intern, Web Search and Mining Group* July 2013 - Dec. 2013
- Advisor: [Zhongyuan Wang](#)
  - **Concept Level Semantic Analysis**  
Proposed and implemented semantic level text comparison based on concept level information extracted from corpus and probabilistic taxonomy.
- Fudan University**, Shanghai, China  
*Research Assistant, Graph Data Management Lab* Oct. 2012 - June 2014
- Advisor: [Prof. Yanghua Xiao](#)

- Designed and built a mechanism to extract information from huge amount of crawled web pages.

PUBLICATIONS / WORKING PAPERS	<p><b>Co-training Embeddings of Knowledge Graphs and Entity Descriptions for Cross-lingual Entity Alignment</b> Muhao Chen, <i>Yingtao Tian</i>, Kai-Wei Chang, Steven Skiena, Carlo Zaniolo <i>Under review for Under Review for 2018 International Joint Conference on Artificial Intelligence</i></p> <p><b>Syntax-Directed Variational Autoencoder for Structured Data</b> Hanjun Dai*, <i>Yingtao Tian*</i>, Bo Dai, Steven Skiena, Le Song. (* Equal Contribution) <i>To Appear in the International Conference on Learning Representations (ICLR), 2018.</i> <i>Best Paper Award, NIPS 2017 workshop for Machine Learning for Molecules and Materials.</i></p> <p><b>Embedding-based Relation Prediction for Ontology Population</b> Muhao Chen, <i>Yingtao Tian</i>, Xuelu Chen, Zijun Xue, Carlo Zaniolo <i>To Appear in the 17th SIAM International Conference on Data Mining (SDM), SIAM 2018.</i></p> <p><b>Towards End-End Generation of High-Resolution Images with Generative Adversarial Networks</b> [Online Demo] Yanghua Jin, Jiakai Zhang, Minjun Li, <i>Yingtao Tian</i>, Huachun Zhu <i>Accepted and presented in spotlight, NIPS 2017 workshop for Machine Learning for Creativity and Design.</i></p> <p><b>Multilingual Knowledge Graph Embeddings for Cross-lingual Knowledge Alignment</b> Muhao Chen, <i>Yingtao Tian</i>, Mohan Yang, Carlo Zaniolo. <i>In Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI), IJCAI/AAAI Press 2017.</i></p>
PROFESSIONAL SERVICES	Reviewer, <i>LREC 2018, COLING 2016</i>
PATENTS	<b>Training and/or Utilizing Recurrent Neural Network Model to Determine Subsequent Source(s) for Electronic Resource Interaction</b> <i>Patent in submission, 2017</i>
HONORS AND AWARDS	<p><b>27th place</b>, 35th Annual World Final of the <a href="#">ACM-ICPC</a> (2011)</p> <p><b>Gold Medal</b>, <a href="#">ACM-ICPC</a> Asia Chengdu Regional Contest, 2011</p> <p><b>Championship and Gold medal</b>, <a href="#">ACM-ICPC</a> Asia Amritapri Regional Contest, 2010</p>
SERVICES	<p>Team member and assistant coach, <a href="#">ACM-ICPC</a> team of Fudan University 2010 - 2013</p> <ul style="list-style-type: none"> <li>• Team member in the <a href="#">ACM-ICPC</a> competition</li> <li>• Assistant coach serving for daily training and competitions overseas.</li> </ul> <p>Main technical staff, <a href="#">ACM-ICPC</a> Shanghai Site Contest Oct. 2011</p> <ul style="list-style-type: none"> <li>• I participated in setting up the system for 2011 <a href="#">ACM-ICPC</a> Shanghai</li> <li>• Sites Contest as one of main technical staffs.</li> </ul>
OTHER EXPERIENCE	<p>Contest Level Problem Solver 2009 - Present</p> <p>An expert on problem solving under alias “alantian”.</p> <p>Participated algorithm contests including Olympiad in Informatics, <a href="#">ACM-ICPC</a>. and <a href="#">Google Code Jam</a></p>