GRACE FAN

EDUCATION

SEPT. 2020 - Present	**Northeastern University** PhD Candidate in COMPUTER SCIENCE PhD Advisor: Professor Renée J. Miller Focus Areas: Data Discovery, Unionable Table Search, Data Integration
MAY 2020	**Brown University** B.S. in COMPUTER SCIENCE Advisor: Professor Ugur Cetintemel

PUBLICATIONS

- **Grace Fan**, Roee Shraga, and Renée J. Miller. "Gen-T: Table Reclamation in Data Lakes" to appear in ICDE 2024 [ArXiv]

- **Grace Fan**, Jin Wang, Yuliang Li, Dan Zhang, and Renée J. Miller. "Semantics-aware Dataset Discovery from Data Lakes with Contextualized Column-based Representation Learning." VLDB 2023 [Paper]

- Aamod Khatiwada, **Grace Fan**, Roee Shraga, Zixuan Chen, Wolfgang Gatterbauer, Renée J. Miller, Mirek Riedewald: "SANTOS: Relationship-based Semantic Table Union Search." SIGMOD 2023 [Paper]

- **Grace Fan**, Wenfei Fan, Yuanhao Li, Ping Lu, Chao Tian, Jingren Zhou: "Extending Graph Patterns with Conditions." SIGMOD 2020 [Paper]

- Amir Ilkhechi, Andrew Crotty, Alex Galakatos, Yicong Mao, **Grace Fan**, Xiran Shi, Ugur Çetintemel: "Deep-Squeeze: Deep Semantic Compression for Tabular Data." SIGMOD Conference 2020 [Paper]

- Stephen Ph. Alexander, Eamonn Kelly, Neil Marrion, John A. Peters, Helen E. Benson, Elena Faccenda, Adam J. Pawson, Joanna L. Sharman, Christopher Southan, Peter Buneman, William A. Catterall, John A. Cidlowski, Anthony P. Davenport, Doriano Fabbro, **Grace Fan**, John C. McGrath, Michael Spedding, Jamie A. Davies, "The Concise Guide to Pharmacology 2015/2016." by The British Journal of Pharmacology, Oct 2015 *PMID: 26650438*

- **Grace Fan**, Wenfei Fan, Floris Geerts: "Detecting Errors in Numeric Attributes". WAIM 2014: 125-137. [Paper]

INTERNSHIPS

2023	**Research Intern**
	Microsoft Research, Redmond, WA
	Worked on using LLMs to facilitate data discovery,
	supporting keyword search and returning relevant datasets.
2022	**Research Intern**
	Megagon Labs, Mountain View, CA
	Helped develop and experiment with different research methods for the problem of
	finding unionable tables, later submitting the paper to VLDB 2023
2018	**Systems Analyst Intern**
	Amazon Web Services, Herndon, VA
	Helped customers troubleshoot Linux and Networking problems
	Assisted customers with internal issues in AWS's services
	Created a website to help the team handle customer cases more efficiently.
2017	**Research Intern**
	Huawei, Beijing, China
	Dealt with database systems and query evaluation
	Researched various diagnostic approaches, specifically with the trace facility
2015	**Research Intern**
	University of Edinburgh, Edinburgh, Scotland
	Helped program a new version of "The Concise Guide to Pharmacology"
	Compiled biomedical data and sorted them by various similarities.
	Used Python to convert it into HTML that produced a PDF resource book

TEACHING EXPERIENCE

2021	Computer Science Department at Northeastern University
	Teaching Assistant for Large-Scale Parallel Data Processing
	Taught two lectures, held office hours, and led students in practice-problem sessions
2018	Computer Science Department at Brown University
	Teaching Assistant for Computer Systems
	Helped students understand concepts, plan out their code structure,
	and debug errors and segmentation faults.

COMPUTER LANGUAGES

Python, Java, SQL, C, C++, Go, Scala, LaTeX, React, HTML, Javascript

INTERESTS AND ACTIVITIES

Northeastern University	Member of PhD Women group
Brown University	Mentor in the Women in Computer Science group;
	Member of the Women in Science and Engineering group

Grace Fan