

# Spencer D Schoenberg

spencrr.dev

Bellevue, WA  
me [at] spencrr.dev  
github.com/spencrr  
linkedin.com/in/spencrr  
google scholar

Proven Distributed Systems & Databases Engineer | MS in CS, 2+ Years Experience

## EXPERIENCE

### Database Engineer — Snowflake

OCT 2024 — SEPT 2025

*Designed and implemented regression detection framework to safeguard performance during rollouts of 20+ of new compiler optimizations*

*Developed historical query plan pinning to ensure (with >5 9s) consistent query performance for latency-sensitive workloads*

### Software Engineer Intern — Microsoft

MAY 2024 — AUG 2024

*Detected and enforced least-privilege access across 100M+ Azure Resources increasing company-wide security posture*

MAY 2023 — AUG 2023

*Built automated BCDR solution for Identity division enabling point-in-time backup and restore of centralized document datastore*

MAY 2022 — AUG 2022

*Developed automated deployment orchestration compute function to update database indexing policies following ring topology and cell-based architecture (CBA)*

MAY 2021 — AUG 2021

*Migrated change audit logging of Azure AD metadata to Cosmos DB leveraging partitioning and replication strategies to enable sync and recovery scenarios at daily scale of 10M events*

### Software Engineer Intern — Johnson Controls

JUNE 2020 — JAN 2021

*Developed Python-based firmware image and code signing tool enhancing end-user device integrity*

*Integrated with and improved Jenkins and Docker CI pipeline, accelerating build times and enhancing dev experience*

## EDUCATION

### University of Wisconsin–Madison

Dean's List

MS Computer Science – Machine Learning, Systems

SEPT 2022 — MAY 2024

BS Computer Science, Data Science

SEPT 2019 — MAY 2022

## TEACHING

### Graduate Teaching Assistant — UW–Madison

SEPT 2023 — MAY 2024 — Big Data Systems (CS 544)

*Led hands-on workshops and employed the Socratic method to simplify challenging, realistic data analysis systems*

SEPT 2022 — MAY 2023 — Computer Graphics (CS 559)

*Interactively instructed 100+ students through graphical theory and practice at varying abstraction layers*

## TECHNICAL SKILLS

### Machine Learning

PyTorch (DDP), scikit-learn, Weak Supervision/Label Synthesis, Hyperparameter Optimization

### Systems & Databases

Cosmos DB, Query Optimization, Distributed Systems

### Data Analysis

SQL, Spark, Azure Synapse, Azure Data Explorer, pandas, NumPy, Data Visualization

### Languages

Python, Java, C#/F.NET, JavaScript/TypeScript, C/C++, R

## PUBLICATIONS

### ScriptoriumWS: A Code Generation Assistant for Weak Supervision

Co-Author – DL4C @ ICLR 2023

*Weak Supervision uses multiple noisy LFs to generate labeled datasets with little-to-no labeled data*

*LFs commonly require domain expertise and are expensive to obtain; this paper explores synthesizing LFs with Code-Generation Models*

### AutoWS-Bench-101: Benchmarking Automated Weak Supervision with 100 Labels

Co-Author – NeurIPS 2022

*A framework for evaluating Automated Weak Supervision techniques in challenging settings against zero-shot methods from Foundation Models*

## RESEARCH

### Research Collaborator

OCT 2021 — MAY 2024

*Collaborated with Professor Aws Albarghouthi, Professor Frederic Sala*

*Employed Generative Models to synthesize Labeling Functions (LFs) for use in Weak Supervision*

*Developed Hyperparameter Search Tool for Regex Synthesis Algorithm*

## AFFILIATIONS

### BSA Eagle Scout

### Stanford Computer Science Intensive

### AI@UW – Artificial Intelligence Club