Spencer D Schoenberg

spencrr.dev

me [at] spencrr.dev
github.com/spencrr
linkedin.com/in/spencer-schoenberg
google scholar

EDUCATION

University of Wisconsin–Madison

PhD Computer Science – Machine Learning, Systems

SEPT 2022 – PRESENT

MS Computer Science – Machine Learning, Systems

SEPT 2022 – MAY 2024

University of Wisconsin–Madison

BS Computer Science, Data Science

SEPT 2019 – MAY 2022

GPA 3.7 – 149 Credits

Stanford University

Computer Science Intensive

JUN 2018 – AUG 2018

GPA 3.81 – 11 Credits

EXPERIENCE

Software Engineer Intern — Microsoft

MAY 2023 – AUG 2023

Created Business Continuity and Disaster Recovery (BCDR) solution using automated point-in-time Backup and Restore for central datastore across Identity business unit

MAY 2022 – AUG 2022

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

MAY 2021 – AUG 2021

Migrated change audit logging for Azure AD to Cosmos DB for sync and recovery

Utilized data partitioning and replication consistency

Achieved end-to-end logging in pre-production environment

Software Engineer Intern — Johnson Controls

JUNE 2020 – JAN 2021

Developed Python-based binary image builder and code signing tool

Improved Jenkins and Docker CI DevOps System

Effectively Collaborated in Agile Scrum Team

TEACHING

Graduate Teaching Assistant — UW–Madison

SEPT 2023 – MAY 2024

Big Data Systems (CS 544) instructed by Tyler R. Caraza-Harter, Meenakshi Syamkumar

SEPT 2022 – MAY 2023

Computer Graphics (CS 559) instructed by Professor Michael Gleicher, Professor Eftychios Sifakis

RESEARCH

Research Collaborator

OCT 2021 – PRESENT

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

Competition Organizer

FEB 2023 – SEPT 2023

AutoML Cup – AutoML Conference 2023

Developed and deployed compute infrastructure for seamless training and evaluation

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

TECHNICAL SKILLS

Languages

Python, C#, Java, JavaScript/TypeScript, R, C/C++

Data Analysis & Machine Learning

NumPy, pandas, Matplotlib, scikit-learn, TensorFlow, PyTorch

AWARDS & AFFILIATIONS

UW–Madison Dean's List

BSA Eagle Scout

Stanford Computer Science Intensive

AI@UW – Artificial Intelligence Club