

Shaobo(Ben) Chen

Boston | +1 857-437-2246 | shaobochoen@outlook.com | [LinkedIn](#) | [GitHub](#)

EDUCATION

Northeastern University

Boston, MA

Master of Computer and Information Sciences [Software Programming & Data Analytics]

to May 2026 | GPA: 4.0

Coursework: Web Development | Mobile Application Development | Database Management Systems | Algorithms | Object Oriented Design

Bachelor of Science Degree in Business Administration [Business Management & Entrepreneurship]

to December 2022 | GPA: 3.6

SKILLS

Languages & Frameworks: Python, R, Java, JavaScript (ES6+), Swift, React, Node.js/Express, FastAPI, HTML5/CSS3

Databases & Cloud: MySQL, PostgreSQL, MongoDB, SQLite, ETL, EDA, R Markdown, Firebase (Firestore/Auth), Pandas, tidyverse/dplyr, ggplot2, Tableau, Chart.js, Superset, kableExtra, AWS, GitHub Actions

Tools: VS Code, Xcode, Jupyter, Anaconda, Cursor, Overleaf, Docker, R studio, Google Colab, GitHub Actions, Aiven, Vercel

Soft Skills: Collaboration, Documentation, Problem Solving, Leadership, Teamwork, Empathy, Adaptability, Curiosity

PROFESSIONAL EXPERIENCES

Yidu Cloud Technology [Healthcare Big Data Firm]

Shanghai, China

Data Analysis Intern | Real-World Evidence Department

May-August 2025

- Engineered SQL pipelines for extracting and preprocessing 120k+ oncology patient records, applying diagnosis, pathology, and treatment rules with joins and window functions, which reduced case identification errors by over 20%.
- Automated inclusion/exclusion funnel workflows in R that operationalized clinical trial-like criteria, reducing screening time from ~2 hours to <30 minutes and enabling reproducible weekly cohort refreshes.
- Converted clinical guidelines into structured variables and engineered features (e.g., staging, treatment line, therapeutic class flags) using SQL and R, ensuring analytic consistency across multiple oncology studies.
- Visualized oncology treatment pathways and prescribing trends using R Markdown, delivering descriptive statistics for final reports on 6 marketed drugs and cutting report preparation time by 30%.

Smith & Nephew [Medical Equipment Firm]

Shanghai, China

Human Resource Intern

July-December 2021

- Orchestrated the Fall 2021 campus recruitment program, developing candidate profiling frameworks and managing assessments with Excel and automated email templates, resulting in the hiring of 40+ interns.
- Mapped and streamlined a pre-onboarding workflow pipeline using Figma process mapping and Workday system integration, which reduced average pre-onboarding time from 3 months to 1 month and decreased manual data input errors by 30%.
- Facilitated onboarding sessions for 60+ new hires, creating standardized training materials and checklists through Workday, which improved onboarding survey satisfaction scores from 78% to 92%.

RELEVANT PROJECTS

Tennis Buddy: Backend Web Application [\[Link\]](#)

Boston, MA | September-December 2024

- Developed full-stack tennis court reservation backend platform with Node.js managing 4,500+ users and 60 courts with real-time availability checking and skill-based player matching **[Node.js, Express.js]**
- Engineered authentication API with JWT tokens and bcrypt password hashing, implementing rate limiting (5 auth attempts/15min) and input validation to prevent unauthorized access and brute-force attacks **[REST API, JWT]**
- Developed responsive frontend using modular ES6 JavaScript with 15+ components, implementing async/await API calls, dynamic form handling, and responsive CSS design across 5 interactive pages **[JavaScript ES6, CSS3, UI/UX]**
- Deployed serverless application on Vercel with MongoDB Atlas, implementing CI/CD pipeline and automated deployments, achieving 99.9% uptime with zero-downtime releases **[Vercel, Cloud Deployment]**

Pre-Acquisition Sales Data Unification Project [\[Link\]](#)

Boston, MA | January-May 2025

- Implemented a cloud-hosted MySQL star-schema data mart to integrate disparate music and film SQLite databases, enabling unified reporting for acquisition readiness **[Aiven MySQL, SQLite]**
- Reconciled schema differences across the two business lines by building ERDs, validating revenue attribution, channel definitions, and country keys, reducing mismatches by 45% **[ERD mapping, schema reconciliation]**
- Constructed production-ready ETL workflows in R + SQL processing ~160k records, implementing batch processing (1,000 rows/batch) with error handling to ensure 100% data completeness and system stability **[ETL, R]**
- Transformed manual reporting processes by implementing R Markdown automation with embedded SQL queries, creating reproducible dashboards that reduced report generation time from ~4 hours to <5 minutes **[R Markdown, kableExtra]**

PeePal: Public Toilet Finder iOS Application [\[Link\]](#)

Boston, MA | September-December 2024

- Led a 4-person team to build a multi-screen iOS app in Swift/UIKit, integrating MapKit for geolocation along with ratings and reporting flows, and tested with 100+ restroom entries over 12 weeks **[GitHub Version Control]**
- Architected backend architecture with Firebase Firestore and Firebase Authentication to support geolocation, ratings, and reporting, enabling secure login and real-time synchronization of user data across devices **[Firebase]**
- Prototyped and translated UI/UX in Figma and translated designs into code, creating a two-step search interface that shortened average search time by 40% in user trials **[Figma]**
- Integrated quality-control features (flagging, reporting, moderation pipeline), resulting in a 25% reduction of invalid entries after test iterations **[Swift, UIKit, MapKit]**