

# SHENGJIA YAN

646-204-9230 · sy2703@nyu.edu · <https://yanshengjia.com> · <https://github.com/yanshengjia>

## EDUCATION

---

**New York University (NYU)**, *M.S. in Computer Engineering*, GPA: 3.8/4.0     *New York* 2019.08 - 2020.12

**Southeast University (SEU)**, *B.S. in Computer Science*, GPA: 3.6/4.0     *Nanjing, China* 2013.08 - 2017.06

## SKILLS

---

- **Languages:** Java, Python, C/C++, Typescript, JavaScript, SQL, HTML, CSS, Shell, L<sup>A</sup>T<sub>E</sub>X
- **Tools:** NAWS, Git, MongoDB, MySQL, Docker, Linux
- **Frameworks:** AWS CDK, Tornado, Flask, Django, React, Bootstrap, Qt, PyTorch, Keras, Scikit-Learn

## WORKING EXPERIENCE

---

**Amazon SDE Intern @ Amazon Payment Products**     *New York* 2020.06 - 2020.08

- Used AWS **CDK** to create and manage **NAWS** infrastructure including a **SQS**, a **SNS** and a **Lambda** function to retrieve data from PayStation by onboarding **CloudAuth**
- Fully launched the project to Amazon **production** environment, reduced A203 error by 93% for ARI lookup Chase API and reduced 2-3 high severity tickets for team monthly
- Executed on-point unit test and end-to-end **Hydra** integration test, set up **CloudWatch** monitoring metrics and alarms for new services

**New York University High Speed Networking Lab Research Assistant**     *New York* 2019.09 - 2019.11

- Implemented a multi-process Web crawler in Python to collect over 100GB of data with **Selenium**
- Designed a real-time In-App activity analysis system in **Python** to classify service types for Internet traffic flow

**17zuoye Software Engineer**     *Beijing, China* 2017.06 - 2019.06

- Designed and implemented a **Web** based automatic essay enhancing system which has been brought online and served millions of K-12 students in China
- Implemented real-time asynchronous updates of the frontend UI using **JavaScript**, **HTML**, **AJAX** and **Bootstrap**
- Built the backend service using **Tornado** with **MongoDB** as database and deployed on **AWS**
- Supported more than 50k tasks simultaneously by implementing **multi-process** services in Python
- Responsible for code integration, unit test, pressure test, build and test automation by integrating Gitlab continuous integration tools (CI/CD) with **Docker**

**Southeast University Knowledge Science and Engineering Lab Research Intern**     *2015.05 - 2017.05*

- Carried out data preprocessing using NLP approaches like spaCy to refine and analyze the text datasets (x1.2 speedup)
- Presented and implemented a Random Walk algorithm in Python based on the Probabilistic Graphical Model to perform word-sense disambiguation on Web tables
- Achieved a 6% increase in F1-score compared to the latest published schemes. The result was **published in [1, 2]**

## SELECTED PROJECTS

---

**Web-based Crowdsourcing Data Annotation Platform (Python, JavaScript)**     *2019.01 - 2019.06*

- Designed a **RESTful** crowdsourcing annotation system with **OAuth** service that supported up to 20k simultaneous tasks
- Increased the consistency of annotators by 50% compared with traditional annotation mechanism by designing multiple quality control algorithms
- Achieved strong data security by writing a **cron** job to gather and backup annotation data from **MongoDB** database automatically every half hour

**C-Minus Compiler (Python)**     *2018.05 - 2018.06*

- Implemented the Regular-Expression-to-NFA converter, LR(1) parser and semantic analysis module in **Python**
- Visualized the compiling process by plotting NFA, DFA, GOTO graphs with GraphViz

**Minecraft-like 3D Game (C++)**     *2017.09 - 2018.01*

- Led a team of four to created a Minecraft-like 3D game based on **OpenGL** APIs in **C++**
- Implemented first-person perspective, third-person perspective and a virtual trackball for complex 3D object rotation
- Supported multiple keyboard functions and mouse functions like Click, Drag-and-Drop, Zoom In/Out

## PUBLICATIONS

---

1. “Language to Network: Conditional Parameter Adaptation with Natural Language Descriptions”, In proceedings of the 58th Annual Meeting of the *Association for Computational Linguistics*, ACL 2020: 6994-7007. [pdf][code]
2. “Entity Linking in Web Tables with Multiple Linked Knowledge Bases”, In proceedings of *Semantic Technology: 6th Joint International Conference, JIST 2016*. Springer, Cham, 2016: 239-253 [pdf]
3. “A Method of Entity Linking in Web Tables based on Multiple Linked Knowledge Bases”, Chinese Patent, CN106503148A, 2017