

# Joshua A. Engels

josh.adam.engels@gmail.com | 301-661-4539 | Houston, TX

GitHub: [github.com/JoshEngels](https://github.com/JoshEngels) Portfolio: [joshengels.com](https://joshengels.com) Google Scholar: [\[link\]](#)

---

## EDUCATION

---

### Rice University

August 2018 - December 2021

*Bachelor of Science in Computer Science*

Houston, TX

*Bachelor of Arts in Mathematics*

Graduated summa cum laude (4.00 GPA)

## RESEARCH EXPERIENCE

---

### ThirdAI Corp

August 2021 - Present

*Artificial Intelligence Engineer*

*Houston, TX*

- Created DESSERT, a general algorithm for performing vector-set search with vector-set queries. When applied to the MSMarco passage retrieval task, DESSERT speeds up an existing state of the art method (ColBERT) by 5X.

### Rice University Sketching and Hashing Lab

August 2020 - August 2021

*Undergraduate Researcher, PI: Dr. Anshumali Shrivastava*

*Houston, TX*

- Implemented and benchmarked FLINNG, a high performance C++ near neighbor search algorithm that uses locally sensitive hashing and group testing to achieve up to a 10x query latency speedup against SOTA methods on high dimensional genome, url, and embedding datasets.
- Proved that FLINNG solves the nearest neighbor problem in sublinear time.

### Rice University Human Computer Interaction Lab

January 2019 - May 2020

*Undergraduate Researcher, PI: Dr. Michael Byrne*

*Houston, TX*

- Built a modular framework in ACT-R and Lisp for simulating a wide variety of paper ballot voting strategies.
- Analyzed the resultant simulated voter error rates in R to identify and characterize ballots that cause systematic voting errors.

## WORK EXPERIENCE

---

### ThirdAI Corp

August 2021 - Present

*Artificial Intelligence Engineer*

*Houston, TX*

- Lead engineer on near neighbor search. Implemented production versions of FLASH, MACH, and DESSERT, all of which are currently being used by customers.
- Made core contributions to ThirdAI's internal machine learning engine, including work on its computation DAG (directed acyclic graph), distributed training, and extensive sparsity-based optimizations.
- Built and currently maintaining/overseeing work on ThirdAI's build system, continuous integration, custom RSA based licensing scheme, and automated PyPi package uploads.

### Databricks

May 2021 - August 2021

*Software Engineering Intern*

*Remote*

- Designed and built a Scala service to monitor and benchmark Kafka clusters. Given a target cluster, the service produced and consumed a constant stream of messages, allowing engineers to monitor cluster availability, message durability, and end to end latency.
- Implemented a monitoring and alerting pipeline for the service and then put it into production running against all of Databricks' Kafka clusters.

## Two Sigma

May 2020 - August 2020

*Software Engineering Intern*

*Remote*

- Built Baikal, a proof of concept data storage service that cached high frequency writes into a PostgreSQL database for later ingestion into Two Sigma's file-based custom time series database, while still maintaining real time read consistency (no stale reads).
- Optimized Baikal's merged read operation, bringing 100 MB throughput from 10 minutes to 10 seconds

## TEACHING EXPERIENCE

---

### Rice University Computer Science Department

January 2020 - May 2020

*Algorithmic Thinking (Comp 182) Teaching Assistant*

*Remote*

- Held weekly office hours, led review sessions, and graded exams and projects for Rice's algorithms and discrete mathematics introduction class.

### Rice University Catalyst Eureka Program

September 2018 - May 2019

*Mentor*

*Remote*

- Mentored a high school student working on a year-long research project to investigate characteristics of popular songs.
- Helped the student learn python, write a program to scrape and analyze historical song data, and create a final presentation for a poster symposium.

## SELECTED HONORS

---

Louis J. Walsh Scholarship in Engineering

2020, 2021

Top 0.18% of ~1 million solvers on ProjectEuler.net

Present

1st place underclassman Rice coding challenge

2018, 2019

National Merit Finalist Scholarship

2018, 2019

## PUBLICATIONS

---

[Practical Near Neighbor Search via Group Testing.](#)

Joshua Engels\*, Benjamin Coleman\*, and Anshumali Shrivastava

NeurIPS 2021: Spotlight talk - top 3%

\* indicates equal contribution

[Missed one! How ballot layout and visual task strategy can interact to produce voting errors.](#)

Joshua Engels, Xianni Wang, Michael D. Byrne.

International Conference on Cognitive Modeling 2020.

[Superintelligence: Consequences of an Intelligence Explosion.](#)

Joshua Engels

Catalyst Rice Undergraduate Science Research Journal (Vol. 12).

## UNDER SUBMISSION

---

[DESSERT: An Efficient Algorithm for Vector Set Search with Vector Set Queries.](#)

Joshua Engels, Benjamin Coleman, Vihan Lakshman, and Anshumali Shrivastava.

## CONFERENCE PRESENTATIONS

---

[How to Deal with the Volume and Velocity Associated with Hundreds of Terabytes \(and Beyond\) of Genomics Data](#)

Ken Kennedy AI and Data Science Conference 2021.

[Practical Near Neighbor Search via Group Testing](#)

NeurIPS 2021: Spotlight talk.

[Missed one! How ballot layout and visual task strategy can interact to produce voting errors.](#)

International Conference on Cognitive Modeling 2020.