Joshua A. Engels

jengels@mit.edu | 301-661-4539 | Cambridge, MA

GitHub: github.com/JoshEngels **Website:** joshengels.com **Google Scholar:** [link]

EDUCATION

Massachusetts Institute of Technology　　　　　　　**August 2023 - August 2028**
PhD in Computer Science　　　　　　　　　　　　　　　　　　　　Cambridge, MA
Advisor: Prof. Max Tegmark

Rice University　　　　　　　　　　　　　　**August 2018 - December 2021**
Bacheler of Science in Computer Science　　　　　　　　　　　　　　　Houston, TX
Bachelor of Arts in Mathematics
Graduated summa cum laude (4.00 GPA)

PUBLICATIONS

Not All Language Model Features are Linear
Joshua Engels*, Isaac Liao*, Eric J. Michaud, Wes Gurnee, and Max Tegmark
Under submission.
* indicates equal contribution

Approximate Nearest Neighbor Search with Window Filters
Joshua Engels, Benjamin Landrum, Shangdi Yu, Laxman Dhulipala, Julian Shun
To appear at the International Conference on Machine Learning (ICML) 2024.

PECANN: Parallel Efficient Clustering with Graph-Based Approximate Nearest Neighbor Search
Shangdi Yu, **Joshua Engels**, Yihao Huang, Julian Shun
Under submission.

DESSERT: An Efficient Algorithm for Vector Set Search with Vector Set Queries
Joshua Engels, Benjamin Coleman, Vihan Lakshman, and Anshumali Shrivastava
Advances in Neural Information Processing Systems (NeurIPS) 2023.

BOLT: An Automated Deep Learning Framework for Training and Deploying Large-Scale Search and Recommendation Models on Commodity CPU Hardware
Nicholas Meisburger, Vihan Lakshman, Benito Geordie, **Joshua Engels**, David Torres Ramos, Pratik Pranav, Benjamin Coleman, Benjamin Meisburger, Shubh Gupta, Yashwanth Adunukota, Siddharth Jain, Tharun Medini, Anshumali Shrivastava
ACM International Conference on Information and Knowledge Management (CIKM) 2023

From Research to Production: Towards Scalable and Sustainable Neural Recommendation Models on Commodity CPU Hardware
Anshumali Shrivastava, Vihan Lakshman, Tharun Medini, Nicholas Meisburger, **Joshua Engels**, David Torres Ramos, Benito Geordie, Pratik Pranav, Shubh Gupta, Yashwanth Adunukota, Siddharth Jain
ACM Conference on Recommender Systems (RecSys) 2023

Practical Near Neighbor Search via Group Testing.
Joshua Engels*, Benjamin Coleman*, and Anshumali Shrivastava
Advances in Neural Information Processing Systems (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

CONFERENCE PRESENTATIONS

DESSERT: An Efficient Algorithm for Vector Set Search with Vector Set Queries
NeurIPS 2023: Poster talk.

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.

SELECTED HONORS

NSF Graduate Research Fellowship (GRFP)	**2024 - 2027**
Louis J. Walsh Scholarship in Engineering	**2020, 2021**
1st place underclassman Rice coding challenge	**2018, 2019**
National Merit Finalist Scholarship	**2018 - 2019**

RESEARCH EXPERIENCE

ThirdAI Corp   **August 2021 - May 2023**
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: Prof. Anshumali Shrivastava   *Houston, TX*

Implemented and benchmarked FLINNG, a high performance C++ nearest neighbor search algorithm that uses locality 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.

WORK EXPERIENCE

ThirdAI Corp   **August 2021 - May 2023**
Artificial Intelligence Engineer   *Houston, TX*

• Lead engineer on nearest neighbor search.

• Made core contributions to ThirdAI's internal machine learning engine, including work on the computation DAG (directed acyclic graph), distributed training, and extensive sparsity-based optimizations.

• Built and maintained ThirdAI's build system, continuous integration, custom RSA based licensing scheme, and automated PyPi package uploads.

TEACHING AND MENTORING EXPERIENCE

MIT Graduate Application Assistance Program   **August 2023 - December 2023**
Mentor   *Cambridge, MA*

Mentored students from underrepresented backgrounds applying to graduate schools.

Rice University Computer Science Department   **January 2020 - May 2020**
Algorithmic Thinking (Comp 182) Teaching Assistant   *Houston, TX*

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   *Houston, TX*

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.

CODING LANGUAGES

Proficient in Python and C++.