Thomas Dybdahl Ahle

San Francisco, CA 94117

thomasahle.com

January 2024

I'm a computer scientist with expertise in algorithms, machine learning and software engineering. I did my PhD in Locality Sensitive Hashing in Copenhagen (2019), while also co-founding an award-winning NLP startup, 2018-19, as Chief of ML. As a Research Scientist at Meta, I focus on scaling language models and ranking.

Education

IT University of Copenhagen

- Ph.D. in Computer Science, June 2019.
- Advisors: Rasmus Pagh, Erik Price.
- Thesis: *Algorithms for Sets and other High Dimensional Data.*

IT University of Copenhagen & University of Copenhagen

- Master of Science in Computer Science, 2017.
- Advisors: Rasmus Pagh, Mikkel Thorup.
- Thesis: *Minhash without false negatives.*

University of Oxford (top 1st)

- Bachelor of Arts in Computer Science, 2013.
- Master of Arts in Computer Science, 2019.
- Advisors: Oege de Moor, Samson Abramsky.
- Thesis: *The Donkey and the Monad.*

Industry

Research Scientist at Meta, 2020 - Present.
I lead and co-founded the Machine Learning Efficiency group, a five-person internal applied research group focused on scaling AI across the organization. During the first year, we developed a new hashing-based algorithm that reduced the size of the internal recommendation systems by 50%. In another project, we devised a new inference algorithm for Bayesian Neural Networks, allowing the Integrity team to deploy well-calibrated models directly on customer devices. Finally, we worked on scaling transformer models, lowering the memory requirement and inference time of Facebook Assistant through the use of embedding table compression, smart attention, and a variety of other tricks.

Chief Machine Learning Officer at SupWiz, 2017 - 2018.
I co-founded an NLP start-up with University of Copenhagen academics. At SupWiz, I lead a four-person team in developing our chatbot software and putting it into production at three of Denmark's largest IT firms. (There are now many more.) We used a combination of traditional symbolic AI and modern (at the time) sentence embeddings. The chatbot was awarded the most prestigious prize by the Innovation Fund Denmark in 2019. I was also in charge of our hiring efforts, interviewing dozens of candidates and hiring four engineers over the course of five months.

Software Engineer at Other: Sophion, Palantir & XION, 2010 - 2014.
Through various software engineering jobs, I have gained broad exposure to the different areas of software development. At Sophion Bioscience, I developed internal debugging tools for sifting through gigabytes of data/second on ion channel screening machines. At Palantir, I ported the Metropolis ontological time-series system (now Foundry) to the web, designing data visualization and efficient processing pipelines. At XION, I Developed the most popular Danish TV listings app for Android at the time, based on data scraped (consensually) from hundreds of TV-station websites.

Awards and Scholarships

Research Travel Award, Stibo-Foundation, 2016.
Highly competitive scholarship, given to just two Danish students a year to collaborate in research abroad.

Northwestern Europe Regional Programming Contest, 1st, Association for Computing Machinery, 2014.
As a member of Lambdabamserne, my team and I made history by becoming the first ever Danish team to qualify for the ACM world finals after winning 1st place out of all universities in North Western Europe.

Danish National Programming Champion, 1st, Netcompany, 2013, 2014.
Crowned Danish National Programming Champion twice, taking the top spot in the Algorithm competition known as "DM i Programmering"

Oxford Computer Science Competition, 1st, University of Oxford, 2013.
For my Numberlink solving software, giving the first fixed parameter polynomial algorithm for the problem.

Demyship, Magdalen College, 2010, 2011.
A historic scholarship awarded to the top students each year.

Les Trophées du Libre, 1st, Free Software Foundation Europe, 2007.
For my work on the PyChess free software chess suite.

Selected Publications

"Thermodynamic Linear Algebra". By M Aifer, K Donatella, M H Gordon, Thomas Dybdahl Ahle, D Simpson, G E Crooks, P J Coles – Submitted 2023.

"Clustering the Sketch: Dynamic Compression for Embedding Tables". By H Tsang, Thomas Dybdahl Ahle at Advances in Neural Information Processing Systems (NeurIPS) (NEURIPS) 2023.

"Tiling with Squares and Packing Dominos in Polynomial Time". By A Aamand, M Abrahamsen, Thomas Dybdahl Ahle, P Rasmussen at Symposium on Computational Geometry (SOCG) 2022.

"Sharp and Simple Bounds for the raw Moments of the Binomial and Poisson Distributions". By Thomas Dybdahl Ahle at Statistics & Probability Letters (STAT. PROB. LETT.) 2022.

"Similarity Search with Tensor Core Units". By Thomas Dybdahl Ahle, F Silvestri at International Conference on Similarity Search and Applications (SISAP) 2020.

"On the Problem of p_1^{-1} in Locality-Sensitive Hashing". By Thomas Dybdahl Ahle at International Conference on Similarity Search and Applications (SISAP) 2020.

"Subsets and Supermajorities: Optimal Hashing-based Set Similarity Search". By Thomas Dybdahl Ahle, J Knudsen at IEEE Symposium on Foundations of Computer Science (FOCS) 2020.

"Oblivious Sketching of High-Degree Polynomial Kernels". By Thomas Dybdahl Ahle, M Kapralov, J Knudsen, R Pagh, A Velingker, D Woodruff, A Zandieh at ACM-SIAM Symposium on Discrete Algorithms (SODA) 2020.

"Optimal Las Vegas Locality Sensitive Data Structures". By Thomas Dybdahl Ahle at IEEE Symposium on Foundations of Computer Science (FOCS) 2017.

"Parameter-free Locality-Sensitive Hashing for Spherical Range Reporting". By Thomas Dybdahl Ahle, M Aumüller, R Pagh at ACM-SIAM Symposium on Discrete Algorithms (SODA) 2017.

"On the Complexity of Inner Product Similarity Join". By Thomas Dybdahl Ahle, R Pagh, I Razenshteyn, F Silvestri at ACM Symposium on Principles of Database Systems (PODS) 2016.

Media

Jon Lund. "En ulv i fåreklæder", Prosa, May 2021. I was interviewed by the IT workers union on Google's FLoC system and the use of SimHash.

"The Stibo-Foundation supports IT-talents", Stibo, August 2016. The announcement of my winning the Stibo Travel grant.

Bidwell, Jonni. "Python: Sunfish chess engine", Linux Format, January 2016. Article about my Sunfish chess software.

"The National Team at the Programming World Cup", Computerworld, June 2015. Coverage of my teams participation in the ICPC World Finals.

Elkær, Mads. "Denmark's Three Greatest Programmers", Computerworld, October 2013.

Open Source Projects

Project Owner at PyChess, 2006 - current.
I developed this chess engine and client for Linux desktop, which became the most pupular way to play chess on the Free Internet Chess Server. Through the years I have lead a team of 4-8 developers and designers, as well as numerous other contributers. Such as the volunteers who translated it to more than 35 languages. In 2009 we won Les Trophées du Libre in Paris.

Project Owner at Sunfish, 2012 - current.
A 111 line python chess engine, which is nevertheless 2000+ rating on the online Lichess server. Because of the simplicity and focus on teaching good AI techniques, it has become a popular project on Github with 2400+ stars and nearly 500 forks. Sunfish was referenced in multiple early applications of neural networks to chess.

Contact

Email: thomas@ahle.dk.

DBLP List of papers

Google Scholar List of Papers

Linkedin: linkedin.com/in/thomasahle

Github: github.com/thomasahle

Twitter twitter.com/thomasahle