

# Anil Shanbhag

## Curriculum Vitae

☎ (+1) 857.253.8697  
✉ [anilshanbhag@gmail.com](mailto:anilshanbhag@gmail.com)  
🌐 <https://anilshanbhag.in/>

### Education

- 2014 - 2020 **Ph.D.**, *Electrical Engineering and Computer Science*,  
Massachusetts Institute of Technology, Cambridge, MA  
Thesis: Interactive Data Analytics using GPUs  
Advisor: Prof. Samuel Madden
- 2014 - 2016 **S.M.**, *Electrical Engineering and Computer Science*,  
Massachusetts Institute of Technology, Cambridge, MA  
Thesis: An Adaptive Partitioning Scheme for Ad-Hoc and Time-Varying Database Analytics  
Advisor: Prof. Samuel Madden
- 2010 - 2014 **B.Tech. (Honors)**, *Computer Science and Engineering*,  
Indian Institute of Technology Bombay, India

### Research Interests

Data Analytics, Query Processing, Performance Engineering

### Work Experience

- Oct, 2020 - **Technical Lead**, *Instabase*, Menlo Park  
Present Instabase helps companies automate their document processing workflows. Lead the Flow Execution team, responsible for managing Instabase's document-processing backend, Flow.  
Developed *Flow V3* - a distributed dataflow system, designed for document processing. The system can execute arbitrary DAGs of UDF functions.  
Oversee key aspects of the system including scalability, reliability, and performance optimization. Flow today processes millions of documents daily for the world's largest financial institutions.  
Collaborate cross-functionally to integrate Flow seamlessly into Instabase's broader platform, enhancing overall automation capabilities for complex document processing workflows.
- May, 2015 - **Software Engineer Intern**, *Ads Infrastructure*, Google, Mountain View, CA  
August, 2015 Mentor: Ben Handy  
Worked on implementing dynamic memory reservation to avoid disk spillage of queries in F1. F1 is the database used to store all the ads data at Google.
- May, 2014 - **Research Intern**, Microsoft Research, Redmond, WA  
August, 2014 Mentor: Dr. Srikanth Kandula  
May, 2013 - Worked on application of sampling-based techniques for *Approximate Query Processing* in *Big Data* context. Incorporated sampling as a first-class operator in SCOPE (Cascades-based optimizer).  
July, 2013 Designed new sampling techniques, created transformation rules for commuting sampling operator with other operators.

### Patent

- [1] **Controlling Approximations of Queries**, Srikanth Kandula, Surajit Chaudhuri, Bolin Ding, *Anil Shanbhag*, Aleksandar Vitorović, Matthaios Olma, Robert Grandl, US Patent 10496643, 2019

### Publications

- [1] **A Study of the Fundamental Performance Characteristics of GPUs and CPUs for Database Analytics**, *Anil Shanbhag*, Samuel Madden, Xiangyao Yu, In *Proceedings of the International Conference on Management of Data (SIGMOD)*, 2020
- [2] **Large-Scale In-Memory Analytics on Intel Optane DC Persistent Memory**, *Anil Shanbhag*, Nesime Tatbul, David Cohen, Samuel Madden, In *Proceedings of Workshop on Data Management on New Hardware (DaMoN)*, 2020

- [3] **Efficient Top-K Query Processing on Massively Parallel Hardware** , *Anil Shanbhag*, Holger Pirk, Samuel Madden, In *Proceedings of the International Conference on Management of Data (SIGMOD)*, 2018
- [4] **Evaluating End-To-End Optimization for Data Analytics Applications in Weld** , Shoumik Palkar, James Thomas, Deepak Narayanan, Pratiksha Thaker, Rahul Palamuttam, Parimajan Negi, *Anil Shanbhag*, Malte Schwarzkopf, Holger Pirk, Saman Amarasinghe, Samuel Madden, Matei Zaharia, *Proceedings of the VLDB Endowment (PVLDB)*, 2018
- [5] **A Robust Partitioning Scheme for Ad-Hoc Query Workloads** , *Anil Shanbhag*, Alekh Jindal, Samuel Madden, Jorge Quiane, Aaron J Elmore, In *Proceedings of the Symposium on Cloud Computing (SoCC)*, 2017
- [6] **Weld: A Common Runtime for High Performance Data Analytics** , Shoumik Palkar, James Thomas, *Anil Shanbhag*, Deepak Narayanan, Holger Pirk, Malte Schwarzkopf, Saman Amarasinghe, Matei Zaharia, *Conference on Innovative Data Systems Research (CIDR)* , 2017
- [7] **Locality-Adaptive Parallel Hash Joins using Hardware Transactional Memory** , *Anil Shanbhag*, Holger Pirk, Sam Madden, *International Workshop on Accelerating Data Analysis and Data Management Systems*, 2016
- [8] **Quickr: Lazily Approximating Complex Adhoc Queries in Bigdata Clusters**, Srikanth Kandula, *Anil Shanbhag*, Aleksandar Vitorovic, Matthaios Olma, Robert Grandl, Surajit Chaudhuri, Bolin Ding , In *Proceedings of International Conference on Management of Data (SIGMOD)*, 2016
- [9] **Optimizing Join Enumeration in Transformation-Based Query Optimizers**, *Anil Shanbhag*, S Sudarshan, *Proceedings of the VLDB Endowment (PVLDB)*, 2014

---

## Skills

Programming C/C++, Python, Javascript, Java, GoLang

---

## Teaching

Fall 2017 Teaching Assistant, 6.830: Database Systems, MIT  
 2012-2013 Manager, Web and Coding Club, IIT-B  
 Fall 2013 Teaching Assistant, CS 305: Computer Architecture, IIT-B  
 Fall 2012 Teaching Assistant, PH 105: Modern Physics, IIT-B

---

## Miscellaneous

2018 - **Creator, Dictanote.co**  
 Ongoing The Dictanote app lets journalists, lawyers, transcriptionists, etc. enhance their productivity by using their voice to type notes. Voice In Chrome Extension lets users use speech to text to dictate on any website like type emails in Gmail or write blog posts on Wordpress. Combined the two applications have over **400k** monthly active users.

2016 Was one of 8 board members of Sangam, the Indian students' association at MIT which has over 800 participants in events throughout the year.

2013 Awarded **Institute Technical Color** at IIT-Bombay for outstanding work in the tech genre; awarded to top 7 out of 8000 total students

2011 Improved border corner rendering and interaction of dashed border style with other border styles in Mozilla Firefox as part of Google Summer of Code