

Alekh Jindal

MIT CSAIL
32 Vassar St., Room 32-G908
Cambridge, MA 02141, USA

Phone: +1 617-710-0572
Email: alekh@csail.mit.edu
Web: <http://people.csail.mit.edu/alekh>

EDUCATION

Massachusetts Institute of Technology (MIT), Cambridge, USA

Postdoctoral Associate (April 2013-present), BigData at CSAIL, Mentor: Samuel Madden

Saarland University & Max Plank Institute for Informatics, Saarbrücken, Germany

Ph.D., *Summa Cum Laude* (2010-2012), Computer Science, Advisor: Jens Dittrich

- Thesis: “OctopusDB: Flexible and Scalable Storage Management for Arbitrary Database Engines”

M.Sc., *Honors* (2008-2010), Computer Science, Advisors: Jens Dittrich & Gerhard Weikum

- Thesis: “Quality in Phrase Mining”

Indian Institute of Technology (IIT), Kanpur, India

B.Tech., (2002-2006) Electrical Engineering, Advisor: Shyama P Das

- Project: “Micro-controller based Power Distribution Monitoring & Control”

RESEARCH INTERESTS

Databases, Large-scale Data-intensive Systems, MapReduce, Graph Analytics, Data Preparation, Data Cleaning, and Physical Database Design.

AWARDS AND HONORS

- VLDB 2014’s *Best Paper Award* for “The Uncracked Pieces in Database Cracking”, 2014
- *Summa Cum Laude* for Doctoral degree. Awarded if “all reporters have evaluated the thesis with this grade and all members of the examination board vote in favour of this”. Committee: Sebastian Hack, Jens Dittrich, Gerhard Weikum, Anastasia Ailamaki, 2012
- CIDR 2011’s *Best Outrageous Ideas and Vision Paper Award* for “Towards a one-size-fits-all Database Architecture”, 2011
- *Strategic Innovation Fund* support for PhD research from Multimodal Computing and Interaction Cluster of Excellence (M2CI), Saarbruecken, Germany, 2010-2012
- *Honors degree* in Computer Science at Saarland University for “finishing studies in a time below the standard duration and the overall grade above 1.3” (scale 1 to 6, with 1 being the best), 2010
- Saarland University’s merit-based tuition fee exemption award for winter term 2009/2010
- Max-Planck Institute for Informatics *IMPRS scholarship* for Master studies. This was awarded to 15 people from 12 different countries across the world, 2008-2010
- British Telecom’s *Iteration Champion* award for best performance in platform development team at Bangalore, India, 2007
- British Telecom’s *Hot-House* winner for best fresher project at Bangalore, India, 2006
- IIT Kanpur’s top-3 final year B.Tech Project in Electrical Engineering, India, 2006
- IIT Joint Entrance Examination’s All India Rank 216 (out of over 160,000 participants), 2002
- 1st in College in Indian School Certificate Examination, 2001
- 3rd in College in Indian Certificate for Secondary Examination, 1999

GRANTS EXPERIENCE

- Served on the National Science Foundation (NSF) Panel on Database Analytics, 2014.
- Helped Prof. Dittrich in writing a 3-year grant to the German Ministry of Education and Science. The proposal received 1.1 million Euros for 6 PhD positions, 2011.

Large-scale Data-intensive Systems

- Database Techniques for Cluster & Cloud Computing (VLDB'10, SOCC'11, VLDB'12, SIGMOD'13)
 - *Integrating database concepts, such as data layouts, indexing, and join processing, into large-scale data flow systems to significantly improve their performance over several workloads.*
 - *Heterogenous data replication in storage systems, with each replica specialized for a different subset of the workload, while still preserving fault-tolerance.*
 - *Performance modeling and resource prediction for database-as-a-service.*
- Data Preparation and Cleaning (SIGMOD'13, under-submissions)
 - *Fine-grained data transformation, including sampling, replication, partitioning, erasure coding, compression, and data placement.*
 - *Query interfaces, join algorithms, and efficient implementation of scalable violation detection and repair.*

Database Architectures & Database Design

- Towards One-size-fits-all Database (VLDB PhD Workshop'10, CIDR'11, CIDR'13)
 - *Building a flexible data management system for the end-users which mimics a variety of specialized systems and adapts to dynamic query workloads. Dropping the assumption of a fixed data store to improve performance and lower costs.*
- Mixed Query Workloads (CIDR'13, VLDB'13, VLDB'14, IEEE BigData'14, under-submission)
 - *Supporting several query workloads in relational databases, including:*
 - (i) *transactional (OLTP) and analytical (OLAP),*
 - (ii) *relational (OLAP) and graph analytics.*
 - (iii) *full-scans and ad-hoc range queries.*
- Database Adaptivity & Robustness (BIRTE'11, VLDB'14, CIDR'15)
 - *Adaptive vertical partitioning to hit the middle ground between row and column layouts.*
 - *Adaptive indexing with faster convergence, robustness, and awareness to modern hardware.*
 - *Data-driven robust data transformations for ad-hoc query workloads.*

[P15] **Alekh Jindal**, Samuel Madden, “GraphiQL: A Graph Intuitive Query Language for Relational Databases” **IEEE BigData**, 441-450, 2014. Acceptance rate: 18.5%

[P14] Felix Martin Schuhknecht, **Alekh Jindal**, Jens Dittrich, “The Uncracked Pieces in Database Cracking” **PVLDB**, 7(2):97-108, 2014. [Best Paper Award]

[P13] **Alekh Jindal**, Endre Palatinus, Vladimir Pavlov, Jens Dittrich, “A Comparison of Knives for Bread Slicing” **PVLDB**, 6(6):361-372, 2013.

[P12] Barzan Mozafari, Carlo Curino, **Alekh Jindal**, Samuel Madden, Performance and Resource Modeling in Highly-Concurrent OLTP Workloads **SIGMOD**, 301-312, 2013.

[P11] **Alekh Jindal**, Jorge-Arnulfo Quiane-Ruiz, Jens Dittrich, “WWHow! Freeing Data Storage from Cages” **CIDR**, 2013.

[P10] **Alekh Jindal**, Felix Martin Schuhknecht, Jens Dittrich, Karen Khachatryan, Alexander Bunte, “How Achaeans Would Construct Columns in Troy” **CIDR**, 2013.

[P9] Jens Dittrich, Jorge-Arnulfo Quiané-Ruiz, Stefan Richter, Stefan Schuh, **Alekh Jindal**, Jörg Schad, “Only Aggressive Elephants are Fast Elephants” **PVLDB**, 5(9):1591-1602, 2012.

[P8] **Alekh Jindal**, Jorge-Arnulfo Quiane-Ruiz, Jens Dittrich, “Trojan Data Layouts: Right Shoes for a Running Elephant” **ACM SOCC**, 21:1-21:14, 2011.

[P7] Jens Dittrich, **Alekh Jindal**, “Towards a one-size-fits-all Database Architecture” **CIDR**, 195-198, 2011. [**Best Outrageous Ideas and Vision Paper Award**]

[P6] Jens Dittrich, Jorge-Arnulfo Quiané-Ruiz, **Alekh Jindal**, Yagiz Kargin, Vinay Setty, and Jörg Schad, “Hadoop++: Making a Yellow Elephant Run Like a Cheetah (Without It Even Noticing)” **PVLDB**, 3(1):518-529, 2010.

OTHER
PUBLICATIONS

[P5] **Alekh Jindal**, “Robust Data Transformations” **CIDR**, 2015. [Abstract]

[P4] **Alekh Jindal**, Praynaa Rawlani, Eugene Wu, Samuel Madden, Amol Deshpande, Mike Stonebraker, “Vertexica: Your Relational Friend for Graph Analytics!” **PVLDB**, 7(13):1669-1672, 2014. [Demo]

[P3] **Alekh Jindal**, Jorge-Arnulfo Quiane-Ruiz, Samuel Madden “Cartilage: Adding Flexibility to the Hadoop Skeleton” **SIGMOD**, 1057-1060, 2013. [Demo]

[P2] **Alekh Jindal**, Jens Dittrich “Relax and Let the Database do the Partitioning Online” **VLDB BIRTE**, 65-80, 2011.

[P1] **Alekh Jindal** “The Mimicking Octopus: Towards a one-size-fits-all Database Architecture” **VLDB PhD Workshop**, 78-83, 2010.

UNDER SUBMIS-
SION/PREPARATION

[U3] Felix Martin Schuhknecht, **Alekh Jindal**, Jens Dittrich, “An Experimental Evaluation and Analysis of Database Cracking”. (Extended Version of [P14]).

[U2] **Alekh Jindal**, Jorge Quiané-Ruiz, Samuel Madden, “Fine-grained Data Preparation and Storage”, Dec 2014.

[U1] **Alekh Jindal**, Samuel Madden, Malú Castellanos, Meichun Hsu, “Graph Analytics using the Vertica Relational Database”, Nov 2014.
<http://arxiv.org/abs/1412.5263>

MISCELLANEOUS
WRITINGS

- **Alekh Jindal**, “Graph Analytics: The New Use Case for Relational Databases”, blog at Intel Science and Technology Center (ISTC) for Big Data, July 2014.
- **Alekh Jindal**, “Benchmarking Graph Databases”, blog at Intel Science and Technology Center (ISTC) for Big Data, Sept 2013.

INTERNATIONAL
PATENTS

- Jens Dittrich, Jorge-Arnulfo Quiané-Ruiz, Stefan Richter, Stefan Schuh, **Alekh Jindal**, Jörg Schad, “Replicated Data Storage System and Methods”. WO Patent WO2013139379.
- Jens Dittrich, **Alekh Jindal**, “A method for storing and accessing data in a database system”. US Patent US20130226959 A1, European Patent EP2614450A1, WO Patent WO2012032184A1.

TEACHING
EXPERIENCE

- Teaching with Educational Technologies, MIT, IAP 2015
- Teaching Certificate Program, MIT, Summer 2014
- Lab Assistant, From ASCII to Answers, MIT, Fall 2013
- TA, Advanced Information Systems Lab, Germany, Winter 2011
- TA, Advanced Information Systems Lab, Germany, Summer 2011
- TA, NOSQL: Managing Data (almost) without a Database System, Germany, Winter 2010
- TA, Advanced Information Systems Lab: OctopusDB, Germany, Summer 2010

- TA, Database Systems core lecture, Germany, Winter 2009
- Research Associate, National Program on Technology Enhanced Learning, India, 2005-2006

MENTORING

- Praynaa Rawlani, M.Eng. Thesis, Graphs analytics on relational databases.
- Felix Martin Schuhknecht, 1st year PhD, Evaluating and improving database cracking algorithms.
- Endre Palatinus, 1st year PhD, Evaluating vertical partitioning algorithms and their impact.
- Karen Khachatryan, 1st year PhD, Techniques for emulating columns stores in row databases.
- Stefan Chouteau, B.Sc Thesis, Implementing a log-structured main-memory database system.
- Sebastian Wendland, M.Sc Thesis, Implementing column store access layer in PostgreSQL.
- Marco Huester, M.Sc Thesis, Applying database cracking over two-dimensional data.
- Felix Martin Schuhknecht, B.Sc Thesis, Compression schemes over hybrid data layouts.

INDUSTRIAL EXPERIENCE

Qatar Computing Research Institute (QCRI), Doha, Qatar

Visiting Scholar, Data Analytics Team

March 2014

- Worked with QCRI scientists to set up a data cleaning research agenda.

Ibibo Web (Naspers Group), Gurgaon, India

Senior Software Engineer, Search Development Team

2007-2008

- Developed vertical search engines using Lucene/Solr technologies and their front-end using LAMP.
- Built web analytics framework to analyze search engine activity, including user behavior, search engine performance, and search quality, on a daily as well as real-time basis.
- Prototyped a lead-generation system for Mobile Ads Business.

British Telecom (BT), Bangalore, India

Associate Consultant, Platform Development Team

2006-2007

- Worked on building an order fulfillment system for the Mobility division.

IBM Software Labs, Pune, India

Summer Intern, Secure Area Network File System Team

Summer 2005

- Implemented an autonomic logging and recovery infrastructure using network blocked devices.

PROFESSIONAL ACTIVITIES

- PC Member, Special Interest Group on Management of Data (SIGMOD), 2015
- PC Member, Proceedings on Very Large Data Bases (PVLDB), 2015
- PC Member, Demo, Extending Database Technology (EDBT), 2015
- PC Member, Data Engineering meets the Semantic Web (DESWeb) Workshop, 2015
- PC Member, Proceedings on Very Large Data Bases (PVLDB), 2014
- Panelist, National Science Foundation (NSF) panel on Database Analytic, 2014
- Referee, SIGMOD Record Research Surveys, 2014
- Referee, Distributed and Parallel Databases (DAPD), 2013
- Thesis Reviewer, "Adaptive Indexing and Covering in OctopusMK" at Saarland University, 2013
- External Reviewer for several conferences: SIGMOD (2014, 2013), VLDB (2013, 2012, 2010), ICDE (2012, 2013), CIKM (2013), EDBT (2011).

TALKS

- Conference on Innovative Database Research (CIDR), Asilomar, California, January, 2015
- IEEE International Conference on BigData, Washington DC, October, 2014
- IEEE International Conference on BigData, Washington DC, October, 2014
- New England Database Summit (NEDB), MIT, Cambridge, Massachusetts, January, 2014
- Intel Head-quarters, Santa Clara, USA, December, 2013
- Symposium on Cloud Computing, Cascais, Portugal, October, 2011
- VLDB, Singapore, September, 2010

REFERENCES

Prof. Samuel Madden

Massachusetts Institute of Technology
32 Vassar Street
Cambridge, MA 02139, USA
☎ +1 (617) 258-6643
✉ madden@csail.mit.edu

Prof. Jens Dittrich

Saarland University
Campus E1.1, Room 311
66123 Saarbrücken, Germany
☎ +49 (681) 302-70141
✉ jens.dittrich@infosys.uni-saarland.de

Prof. Gerhard Weikum

Max-Planck Institute for Informatics
Campus E1.4, Room 402
66123 Saarbrücken, Germany
☎ +49 (681) 9325-500
✉ weikum@mpi-inf.mpg.de