Christopher Re
Associate Professor of Computer Science

Bio

Christopher (Chris) Ré is an associate professor in the Department of Computer Science at Stanford University in the InfoLab who is affiliated with the Statistical Machine Learning Group, Pervasive Parallelism Lab, and Stanford AI Lab. His work's goal is to enable users and developers to build applications that more deeply understand and exploit data. His contributions span database theory, database systems, and machine learning, and his work has won best paper at a premier venue in each area, respectively, at PODS 2012, SIGMOD 2014, and ICML 2016. In addition, work from his group has been incorporated into major scientific and humanitarian efforts, including the IceCube neutrino detector, PaleoDeepDive and MEMEX in the fight against human trafficking, and into commercial products from major web and enterprise companies. He cofounded a company, based on his research, that was acquired by Apple in 2017. He received a SIGMOD Dissertation Award in 2010, an NSF CAREER Award in 2011, an Alfred P. Sloan Fellowship in 2013, a Moore Data Driven Investigator Award in 2014, the VLDB early Career Award in 2015, the MacArthur Foundation Fellowship in 2015, and an Okawa Research Grant in 2016.

ACADEMIC APPOINTMENTS
• Associate Professor, Computer Science
• Member, Bio-X

PROGRAM AFFILIATIONS
• Stanford SystemX Alliance

LINKS
• My homepage: http://cs.stanford.edu/people/chrismre/

Research & Scholarship

CURRENT RESEARCH AND SCHOLARLY INTERESTS
Algorithms, systems, and theory for the next generation of data processing and data analytics systems.

PROJECTS
• DeepDive - Stanford (7/1/2013 - present)

Teaching

COURSES
2019-20
• Machine Learning: CS 229, STATS 229 (Aut, Spr)

2018-19
• Machine Learning: CS 229, STATS 229 (Spr)
2016-17

• Introduction to Databases: CS 145 (Aut)

STANFORD ADVISEES

Doctoral Dissertation Reader (AC)
Matthew Feldman, Abhijeet Mohapatra, Manolis Papadakis, Tian Zhao

Postdoctoral Faculty Sponsor
Charles Kuang, Sharon Li, Avner May, Laurel Orr, Fred Sala

Doctoral Dissertation Advisor (AC)
Ines Chami, Tri Dao, Beliz Gunel, Nimit Sohoni

Master's Program Advisor
Ramnik Arora, Ashkon Farhangi, Min Kim, Bill Lin, Yawen Shen

Doctoral Dissertation Co-Advisor (AC)
Anand Avati, Steven Diamond, Dan Fu, Sarah Hooper

Doctoral (Program)
Karan Goel, Albert Gu, Bryan He, Megan Leszczynski, Sen Wu, Jian Zhang

Publications

PUBLICATIONS

• Ringtail: Nowcasting Made Easy
  Antenucci, D., Cafarella, M., Levenstein, Margaret, C., Ré, C., Shapiro, M.
  2013

• Building an Entity-Centric Stream Filtering Test Collection for TREC 2012
  2013

• GeoDeepDive: Statistical Inference using Familiar Data-Processing Languages. SIGMOD 13 (demo).
  Zhang, C., Govindaraju, V., Borchardt, J., Foltz, T., Ré, C., Peters, S.
  2013

• Parallel Stochastic Gradient Algorithms for Large-Scale Matrix Completion Mathematical Programming Computation
  Recht, B., Ré, C.
  2013

• Using Commonsense Knowledge to Automatically Create (Noisy) Training Examples from Text StarAI with AAAI
  Natarajan, S., Picado, J., Khot, T., Kersting, K., Ré, C., Shavlik, J.
  2013

• Understanding Tables in Context Using Standard NLP Toolkits ACL 2013 (Short Paper)
  Govindaraju, V., Zhang, C., Ré, C.
  2013

• Hazy: Making it Easier to Build and Maintain Big-data Analytics
  Kumar, A., Niu, F., Ré, C.
  2013

• Towards High-Throughput Gibbs Sampling at Scale: A Study across Storage Managers.
  Zhang, C., Ré, C.
• Robust Statistics in IceCube Initial Muon Reconstruction
  Wellons, M., Collaboration, t., Recht, B., Ré, C.
  2013

• Feature Selection in Enterprise Analytics: A Demonstration using an R-based Data Analytics System
  Konda, P., Kumar, A., Ré, C., Sashikanth, V.
  2013

• An Approximate, Efficient LP Solver for LP Rounding *NIPS*
  Sridhar, S., Bittorf, V., Liu, J., Zhang, C., Ré, C., Wright, Stephen, J.
  2013

• Brainwash: A Data System for Feature Engineering (Vision Track)
  Anderson, M., Antenucci, D., Bittorf, V., Burgess, M., Cafarella, M., Kumar, A.
  2013

• Ringtail: Nowcasting Made Easy.
  Antenucci, D., Li, E., Liu, S., Cafarella, Michael, J., Ré, C.
  2013

• DeepDive: Web-scale Knowledge-base Construction using Statistical Learning and Inference *VLDS*
  Niu, F., Zhang, C., Ré, C., Shavlik, J.
  2012

• Toward a noncommutative arithmetic-geometric mean inequality: conjectures, case-studies, and consequences *COLT*
  Recht, B., Ré, C.
  2012

• Big Data versus the Crowd: Looking for Relationships in All the Right Places *ACL*
  Zhang, C., Niu, F., Ré, C., Shavlik, J.
  2012

• Factoring nonnegative matrices with linear programs. *NIPS*
  Bittorf, V., Recht, B., Ré, C., Tropp, Joel, A.
  2012

• Scaling Inference for Markov Logic via Dual Decomposition (Short Paper). *ICDM*
  Niu, F., Zhang, C., Ré, C., Shavlik, J.
  2012

• Elementary: Large-scale Knowledge-base Construction via Machine Learning and Statistical Inference *IJSWIS, Special Issue on Knowledge Extraction from the Web, 2012, to appear*
  Niu, F., Zhang, C., Ré, C., Shavlik, J.
  2012

• Understanding cardinality estimation using entropy maximization *ACM Trans. Database Syst.*
  Ré, C., Suciu, D.
  2012; 37: 6

• Worst-case Optimal Join Algorithms *PODS*
  Ngo, Hung, Q., Porat, E., Ré, C., Rudra, A.
  2012

• The MADlib Analytics Library or MAD Skills, the SQL. *PVLDB*
  2012

• Probabilistic Management of OCR using an RDBMS
  Kumar, A., Ré, C.
Towards a Unified Architecture for In-Database Analytics
Feng, A., Kumar, A., Recht, B., Ré, C.
2012

Optimizing Statistical Information Extraction Programs Over Evolving Text
Chen, F., Feng, X., Ré, C., Wang, M.
2012

Probabilistic Databases
Suciu, D., Olteanu, D., Ré, C., Koch, C.
2011

Automatic Optimization for MapReduce Programs PVLDB
Jahani, E., Cafarella, Michael, J., Ré, C.
2011; 4: 385-396

Tuffy: Scaling up Statistical Inference in Markov Logic Networks using an RDBMS PVLDB
Niu, F., Ré, C., Doan, A., Shavlik, Jude, W.
2011; 4: 373-384

Dalvi, Nilesh, N., Re, C., Suciu, D.
2011; 77: 473-490

Felix: Scaling Inference for Markov Logic with an Operator-based Approach ArXiv e-prints
Niu, F., Zhang, C., Ré, C.
2011

Hogwild!: A Lock-Free Approach to Parallelizing Stochastic Gradient Descent NIPS
Niu, F., Recht, B., Ré, C., Wright, Stephen, J., Ré, C., Shavlik, J.
2011

Parallel Stochastic Gradient Algorithms for Large-Scale Matrix Completion Optimization Online
Recht, B., Ré, C.
2011

Incrementally maintaining classification using an RDBMS PVLDB
Koc, M. L., Ré, C.
2011; 4: 302-313

Manimal: Relational Optimization for Data-Intensive Programs WebDB
Cafarella, Michael, J., Ré, C.
2010

Approximation Trade-Offs in a Markovian Stream Warehouse: An Empirical Study (Short Paper) ICDE
Letchner, J., Ré, C., Balazinska, M., Philipose, M.
2010

Understanding Cardinality Estimation using Entropy Maximization PODS
Ré, C., Suciu, D.
2010

Transducing Markov Sequences PODS
Kimelfeld, B., Ré, C.
2010

Query Containment of Tier-2 Queries over a Probabilistic Database Management of Uncertain Databases (MUD)
Moore, Katherine, F., Rastogi, V., Ré, C., Suciu, D.
2009

- **Lahar Demonstration: Warehousing Markovian Streams** *PVLDB*
  Letchner, J., Ré, C., Balazinska, M., Philipose, M.
  2009; 2: 1610-1613

- **Large-Scale Deduplication with Constraints Using Dedupalog** *ICDE*
  Arasu, A., Ré, C., Suciu, D.
  2009; 952-963

- **The Trichotomy of HAVING Queries on a Probabilistic Database** *VLDB Journal*
  Ré, C., Suciu, D.
  2009

- **Access Methods for Markovian Streams** *ICDE*
  Letchner, J., Ré, C., Balazinska, M., Philipose, M.
  2009; 246-257

- **Probabilistic databases: Diamonds in the dirt** *Commun. ACM Volume*
  Dalvi, Nilesh, N., Ré, C., Suciu, D.
  2009; 52: 86-94

- **General Database Statistics Using Entropy Maximization** *DBPL*
  Kaushik, R., Ré, C., Suciu, D.
  2009; 84-99

- **Managing Large-Scale Probabilistic Databases** *University of Washington, Seattle*
  Ré, C.
  2009

- **Repeatability & Workability Evaluation of SIGMOD 2009** *SIGMOD Record*
  Manegold, S., Manolescu, I., Afanasiev, L., Feng, J., Gou, G., Hadjieleftheriou, M., Re, C. M.
  2009; 38: 40-43

- **Implementing NOT EXISTS Predicates over a Probabilistic Database** *QDB/MUD*
  Wang, T., Ré, C., Suciu, D.
  2008: 73-86

- **A demonstration of Cascadia through a digital diary application**
  Khoussainova, N., Welbourne, E., Balazinska, M., Borriello, G., Cole, G., Letchner, J.
  2008

- **Managing Probabilistic Data with Mystiq (Plenary Talk)**
  Ré, C.
  2008

- **Systems aspects of probabilistic data management (Part II)** *PVLDB*
  Balazinska, M., Ré, C., Suciu, D.
  2008; 1: 1520-1521

- **Systems aspects of probabilistic data management (Part I)** *PVLDB*
  Balazinska, M., Ré, C., Suciu, D.
  2008; 1: 1520-1521

- **Approximate lineage for probabilistic databases** *PVLDB*
  Ré, C., Suciu, D.
  2008; 1: 797-808

- **Managing Probabilistic Data with MystiQ: The Can-Do, the Could-Do, and the Can’t-Do** *SUM*
  Ré, C., Suciu, D.
• Event queries on correlated probabilistic streams
  Ré, C., Letchner, J., Balazinska, M., Suciu, D.
  2008

• Advances in Processing SQL Queries on Probabilistic Data
  Ré, C., Suciu, D.
  2008

• Challenges for Event Queries over Markovian Streams
  IEEE Internet Computing
  Letchner, J., Ré, C., Balazinska, M., Philipose, M.
  2008; 12: 30-36

• Structured Querying of Web Text Data: A Technical Challenge
  CIDR
  Cafarella, Michael, J., Ré, C., Suciu, D., Etzioni, O.
  2007: 225-234

• Managing Uncertainty in Social Networks
  Adar, E., Ré, C.
  2007; 30: 15-22

• Materialized Views in Probabilistic Databases for Information Exchange and Query Optimization
  VLDB
  Re, C., Suciu, D.
  2007: 51-62

• Efficient Top-k Query Evaluation on Probabilistic Data
  ICDE
  Ré, C., Dalvi, Nilesh, N., Suciu, D.
  2007: 886-895

• Efficient Evaluation of HAVING Queries
  DBPL
  Ré, C., Suciu, D.
  2007: 186-200

• Management of data with uncertainties
  CIKM
  Re, C., Suciu, D.
  2007: 3-8

• Orderings on Annotated Collections
  Liber Amicorum in honor of Jan Paredaens 60th Birthday
  Ré, C., Suciu, D., Tannen, V.
  2007

• A Complete and Efficient Algebraic Compiler for XQuery
  ICDE
  Re, C., Sim'eon, J., Fern'andez, Mary, F.
  2006: 14

• XQuery!: An XML Query Language with Side Effects
  Ghelli, G., Ré, C., Sim'eon, J.
  2006

• Query Evaluation on Probabilistic Databases
  Ré, C., Dalvi, Nilesh, N., Suciu, D.
  2006; 29: 25-31

• MYSTIQ: a system for finding more answers by using probabilities
  Boulos, J., Dalvi, Nilesh, N., Mandhani, B., Mathur, S., Ré, C., Suciu, D.
  2005

• A Framework for XML-Based Integration of Data, Visualization and Analysis in a Biomedical Domain
  XXSym
  Bales, N., Brinkley, J., Lee, E., Sally, Mathur, S., Re, C., Suciu, D.
2005: 207-221

- **Supporting workflow in a course management system** *SIGCSE*
  Botev, C., Chao, H., Chao, T., Cheng, Y., Doyle, R., Grankin, S.
  2005: 262-266

- **Distributed XQuery**
  Ré, C., Brinkley, J., Hinshaw, K., Suciu, D.
  2004

- **WS-Membership - Failure Management in a Web-Services World** *WWW (Alternate Paper Tracks)*
  Vogels, W., Ré, C.
  2003

- **A Collaborative Infrastructure for Scalable and Robust News Delivery**
  Vogels, W., Ré, C., Renesse, R., Birman, Kenneth, P.
  2002